# GazeRoomLock: Using Gaze and Head-pose to Improve the Usability and Observation Resistance of 3D Passwords in Virtual Reality

Ceenu George[1], Daniel Buschek[2], Andrea Ngao[1] and Mohamed Khamis[3]

[1] Chair for Media informatics, LMU Munich, ceenu.george@ifi.lmu.de
[2] Research Group HCI + AI, Department of Computer Science, University of Bayreuth, daniel.buschek@uni-bayreuth.de
[3] School of Computing Science, University of Glasgow, mohamed.khamis@glasgow.ac.uk

**Abstract.** Authentication has become an important component of Immersive Virtual Reality (IVR) applications, such as virtual shopping stores, social networks, and games. Recent work showed that compared to traditional graphical and alphanumeric passwords, a more promising form of passwords for IVR is 3D passwords. This work evaluates four multimodal techniques for entering 3D passwords in IVR that consist of multiple virtual objects selected in succession. Namely, we compare eye gaze and head pose for *pointing*, and dwell time and tactile input for *selection*. A comparison of a) usability in terms of entry time, error rate, and memorability, and b) resistance to real world and offline observations, reveals that: multimodal authentication in IVR by pointing at targets using gaze, and selecting them using a handheld controller significantly improves usability and security compared to the other methods and to prior work. We discuss how the choice of pointing and selection methods impacts the usability and security of 3D passwords in IVR.

## 1 Introduction

Recent advances in immersive virtual reality (IVR) using head mounted displays (HMDs) allow users to shop in virtual stores, visit virtual social networking sites, and experience highly immersive games. These applications demand authentication to confirm users' identity to, for example, perform purchases or log in. At the same time, HMDs are becoming self-contained wireless devices, without external input devices such as keyboards [39]. These trends underline the need for secure and usable authentication that seamlessly integrates into the mobile and ubiquitous IVR experience.

To appear in the 7th International Conference on Augmented Reality, Virtual Reality and Computer Graphics (AVR 2020) – Lecture Notes in Computer Science, LNCS 12242, Springer"

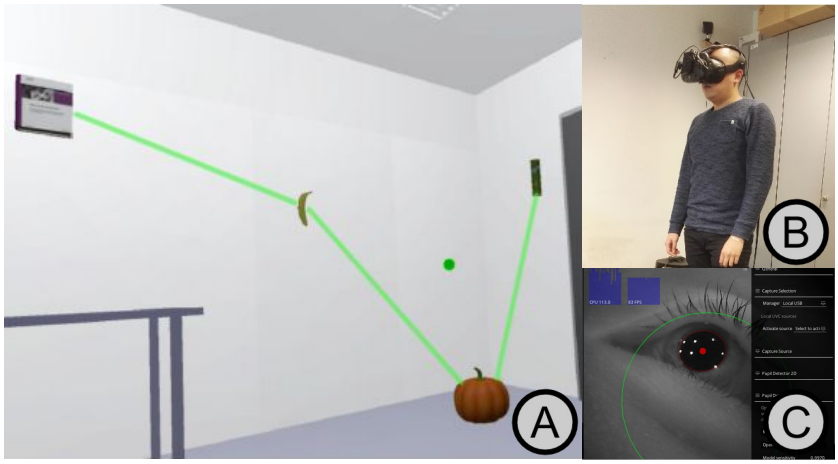


**Fig. 1.** We compare four input techniques for authentication in IVR by selecting a series of 3D targets. MultimodalGaze and MultimodalHead are the fastest, least error-prone, and most secure against real world observations. MultimodalGaze is even more resilient to video observations.

Previous work attempted to transfer authentication concepts from mobile devices to IVR. For example, George et al. [18] experimented with PINs and Android lock patterns in IVR. However, they found that by observing the user during authentication, bystanders in the real world can infer their input [18]. This is a growing threat in the context of IVR with HMDs, in which HMDs are becoming more immersive and users are blindfolded from the real world – making it less likely for users to be aware of bystanders. Additionally, the need to hide IVR users' interactions from bystanders is exacerbated by the affordance of mobile HMDs, such as the Oculus Go [39], which are increasingly used in public settings [19, 35] and can be shared across multiple users in the same household. A recent more promising solution for authentication in IVR is by using the handheld controllers to point at virtual 3D objects that make up the password [16]. While that approach's adoption of 3D passwords made it more suitable for IVR, authentication times were relatively long (between 8.58 s and 14.32 s). They also found that the use of HMD controllers while authenticating is prone to observation attacks [16, 18]. In the context of 3D authentication for IVR, our research questions summarize as follows:

**R1** Do modalities that are hidden from the bystander, such as gaze- and head-based interaction, improve security while maintaining usability?
**R2** How does multimodal interaction impact usability and observation resistance in IVR?

To this end, we compare: 1) **UnimodalGaze**: pointing via gaze and selection via dwell time, 2) **MultimodalGaze**: pointing via gaze and selection via tactile input, 3) **UnimodalHead**: pointing via head-pose and selection via dwell time,

and 4) **MultimodalHead**: pointing via head-pose and selection via tactile input. Our choice of methods was motivated by the advantages of using gaze to support authentication as outlined by previous work [24]. In two studies, we evaluate the techniques' impact on a) **usability** of the scheme (N=48) in terms of entry time, error rate and memorability, and on b) the **observation resistance** of input (N=26) against real time observations, and offline video attacks. Based on our analysis of usability, memorability and security against two realistic threat models, we recommend **MultimodalGaze** for entering 3D passwords in IVR: We found it to be significantly faster (5.94 s) compared to methods in prior work, significantly less error prone (1.32% error rate), and significantly more resilient to real world observations (18% attack success rate) and offline observations using recordings of the user and the virtual scene (10% attack success rate). We discuss how the choice of pointing and selection methods impacts the usability and security of 3D passwords.

## 2 Related Work

**Authentication in Virtual Reality.** 3D virtual environments provide a number of advantages which support the authentication process. Firstly, they offer a limitless space for password creation. Secondly, they aid memorability by utilizing human spatial memory to recall passwords [5]. Alsulaiman et al. [3] proposed an authentication concept, where passwords consist of users' navigation through the virtual space and their interaction with objects; for example, walking to the first room and sitting down. Gurary et al. [20] and George et al. [16] transferred this concept to immersive virtual reality with an HMD. In the latter authentication concept, users could authenticate by pointing at objects in the 3D environment rather than tapping it or interacting with it actively [16]. However, authentication in IVR also has a main drawback, which is users' inability to view the real world while they are interacting with an HMD. In this context, they are unaware of real world bystanders that might be observing their body and/or arm gestures from which they can infer the entered password. Thus, we improve prior work in 3D authentication by utilizing inconspicuous gaze interactions.

**Gaze-supported Authentication.** Early work on gaze-based interaction recognized authentication as one of the main domains that can benefit from the subtle nature of eye gaze. For example, EyePassShapes used gaze gestures for authentication [10], while CGP is a graphical scheme that allowed users to gaze and dwell at certain positions on pictures to authenticate [15]. More recently, CueAuth allowed users to authenticate by gazing at on-screen moving targets [30]. These schemes performed well against observations, but were slow with average authentication times ranging from 12.5 seconds in EyePassShapes [10], 36.7 seconds in CGP [15], and 26.35 seconds in CueAuth [30]. Other works employed gaze for multimodal authentication. Kumar et al. [32] proposed EyePassword, in which users gaze at an on-screen digit and press the space bar to select it. Other multimodal schemes include GazeTouchPIN [28], where users authenticated using gaze and touch. Abdrabou et al. [1] compared multiple techniques that used


**Fig. 2.** Pointing Methods. In GazeRoomLock, the area that is pointed at is indicated by a green dot. When using UnimodalGaze and MultimodalGaze, the point follows the user's eye movements (A). While in UnimodalHead and MultimodalHead the point is centered in the middle of the user's view (C). Objects are highlighted once they are pointed at (B and D).

gaze, mid-air gestures or combinations of both for authentication. Their gaze-based scheme was a replication of EyePassword [32], yet it outperformed the other 5 techniques and even the original version by Kumar et al. The authors underlined the value of replicating prior eye tracking applications in light of the improved hardware and gaze estimation algorithms. The use of several modalities for password entry results splits the observer's attention, which improves security. Multimodal schemes were generally faster than the unimodal counterparts – 9.2 s in EyePassword [32] and 10.8 s in GazeTouchPIN [28]. The aforementioned schemes relied on the knowledge factor. Other schemes leveraged gaze for behavioral biometric authentication [31, 42, 43, 26] and the latest HMDs offer built-in solutions, such as the retina authentication [23, 44]. While behavioral biometrics offers continuous authentication and can be fast, they require sharing personal data with third parties, cannot be changed or invalidated if leaked, and are incompatible with many existing backends.

Our work is unique compared to prior work in that it investigates multimodal approaches to enter 3D passwords in IVR. Results show that our MultimodalGaze and MultimodalHead methods are faster (5.92 s and 5.51 s) compared to prior work (e.g., 9.2 s [32], 10.8 s [28], 12.5 [10], and 26.35 s [30]). Our work significantly improves over a recently proposed authentication scheme for IVR [16]. Namely, our methods leverage the user's eye gaze for authentication in virtual reality. Instead of pointing at targets, users gaze at the target they want to select as part of their password. This improvement resulted in higher usability: lower entry time (5.92 s - 5.51 s vs 8.58 s - 14.32 s [16]), and lower error rate (0.46% - 1.39% vs 2.78% [16]). Our aim is not to propose a new concept but rather to explore input methods for 3D passwords in IVR: We improve entry time and error rates over prior work (e.g., [16]) despite using the same password space.

## 3 Concept

We will refer to our authentication concept as GazeRoomLock. In GazeRoomLock, users authenticate by selecting a number of 3D objects in a virtual room. For example, in Figure 1, the user selects a book, followed by a banana, pumpkin, and then a can of chips. Because the user is wearing an HMD, bystanders

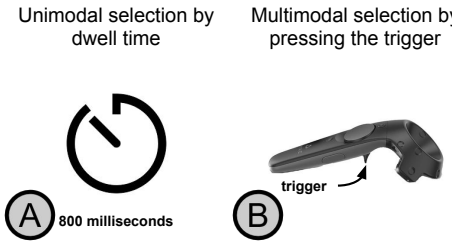


**Fig. 3.** We experimented with two selections methods in GazeRoomLock. Namely, users can select the object that is pointed at by either a) dwelling at it for 800 ms (UnimodalGaze and UnimodalHead) or b) pressing the controller's trigger (MultimodalGaze and MultimodalHead).

cannot see the virtual scene that the user is in[4]. In our current implementation and in our study, 3D passwords are of length 4, and users could choose from 9 selectable objects. While future implementations can allow longer passwords, and integrate more selectable objects, we chose a length of 4 to be comparable to prior work on authentication [18, 16, 30, 47]. Our scheme was influenced by Lock Patterns that are commonly used on Android devices. Lock patterns provide an entropy of 389,112 [6]. Unlike Lock Patterns, our scheme allows selecting the same object multiple times, but not in succession. This means our system has a higher entropy than Lock Patterns.

### 3.1 Pointing and Selecting

The selection of each target in GazeRoomLock comprises of two steps: *Pointing* at a target, and then (2) *Selecting* it. We studied different approaches for achieving each step.

**Pointing** There are two predominant ways for hands-free pointing in virtual reality: 1) users can point using their **eye gaze**, i.e., by looking at the target, or 2) users can adjust their **head pose** such that the target is in the middle of their view. Eye gaze is fast (e.g., faster than pointing [33, 41]), and is likely to be very secure due to its subtleness. On the other hand, while changing the head pose can be more obvious to bystanders, it is accurate for pointing [33], does not induce eye fatigue, and was suggested as a proxy for gaze [4, 14]. In our study, pointing by gaze was achieved by eye tracking. The gaze point was visualized using a green dot (Figure 2A). Pointing by head pose was done by moving a point at the center of the user's view to overlay the target (Figure 2C).

**Selection** After successfully pointing at the target, selection can be done in multiple ways. We experimented with the following designs: 1) In **Unimodal**

[4] Some HMDs show the user's view on a nearby screen. This feature must be automatically disabled during authentication. We expect all HMDs will become untethered.

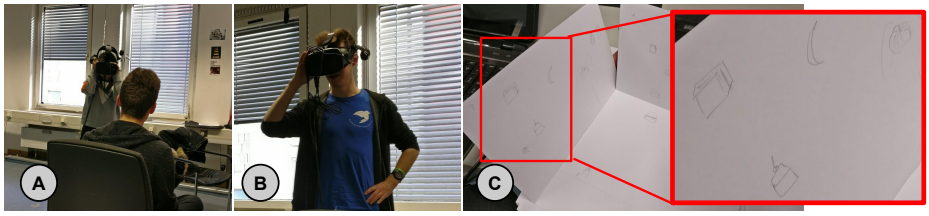


**Fig. 4.** In the security evaluation of GazeRoomLock, participants first observed the experimenter as she entered a GazeRoomLock password in VR (A). Participants then put on the HMD and tried providing up to three guesses of the observed password (B). Some participants came up with creative ways to note down their observations. For example, one of the participants folded two sheets of papers to make a replica of the virtual environment (C).

**selection**, the same modality used for pointing (i.e., eye gaze or head pose) is also used for selection. We realized this using dwell time, which means that users had to fixate at the target for a certain amount of time for the system to distinguish intentional selections from perception of targets. To determine suitable dwell durations, we conducted a pilot study (N=6) in which we experimented with three different dwell time durations that are based on prior work on gaze-based password and text entry [10, 36, 38]: 400 ms, 500 ms, and 800 ms. The results indicate that a dwell time of 800 ms is the least error prone, and also the fastest because participants spend less time recovering from errors. Therefore, we used 800 ms in the Unimodal selection condition. In 2) **Multimodal selection**, an additional modality was used to confirm selection of the target that is pointed at [22]. In this version of GazeRoomLock, we chose to confirm the selection of the pointed at target using the handheld controller's trigger button (Figure 3B). There are other means for selection via eye gaze that leverage gaze behavior rather than fixations. Examples include gaze gestures [12] and Pursuits [13, 45]. While these approaches are suitable for selection in IVR , they require either a) arranging targets in a fixed way [12] which is vulnerable to observations, or b) continuously moving the targets in a predefined manner [13, 14, 29] which requires long authentication times (e.g., 26.35 s [30]).

## 4 Implementation

We implemented GazeRoomLock using Unity 3D with C#. We designed a virtual room that replicated a real room in our lab, and ten selectable 3D objects using Blender. The virtual objects were selected to cover a range of shapes and colors, and to resemble every day objects so users would relate to them. Their sizes matched the size of their real world equivalents. The selectable objects are: A Pringles can, a lemon, a banana, a pumpkin, a spray bottle, a smartphone, a booklet, a plant, a bottle of water, a soda drink can and a trash can. The HTC Vive was equipped with a binocular eye tracker by Pupil Labs [34] to enable gaze-based selection. We used the built-in calibration of the Pupil labs software.

A green dot represents the area that is pointed at. The point is placed where the user is looking in case of gaze-based pointing (Figure 2A), or at the center of the user's view in case of head pose selection (Figure 2C). If this point falls within a predefined area around an object (the "collider"), the object is considered "being pointed at" and is highlighted (see Figures 2B and 2D).

In case of unimodal selection, we employ the concept of dwell time, i.e., pointing initiates a timer which is reset when the object is no longer being pointed at. If an object is pointed at for 800 ms, it is considered selected. Recall that the 800 ms dwell time threshold was chosen based on the pilot study reported in Section 3.1. In case of multimodal selection, an object is considered selected if the user presses the controller's trigger while pointing at the object. Once selected, the object is highlighted (see Figures 2C and 2D). The objects are highlighted until either a complete password has been entered, or the user undoes the last entry by selecting a virtual trash can. A blue line connects consecutively selected objects. The line turns green if the password is correct, and red otherwise. This feedback design is inspired by Android's pattern locks. While GazeRoomLock is inspired by RoomLock by George et al. [16], this is the first implementation of the aforementioned techniques for authentication using 3D passwords in IVR.

## 5 Usability Evaluation

We ran a user study to assess the impact of multiple design factors on GazeRoomLock. Namely, we evaluate how GazeRoomLock's usability and memorability are influenced by the choice of:

- **UnimodalGaze**: Pointing via gaze, and selecting by dwelling.
- **MultimodalGaze**: Pointing via gaze, and selecting by pressing a controller's trigger button.
- **UnimodalHead**: Pointing via head pose, and selecting by dwelling.
- **MultimodalHead**: Pointing via head pose, and selecting by pressing a controller's trigger button.

### 5.1 Study Design

The study was split into two parts: 1) a lab study, where we investigated the usability of GazeRoomLock, and 2) a follow up remotely administered questionnaire to gain insights into password memorability. The lab study followed a mixed-subjects design with two independent variables:

- **Pointing method**, a between-subjects variable with two conditions: Gaze vs Head-pose (see Figure 2). We chose a between-subjects design to avoid potential learning effects which could bias the results.
- **Selection method**, a within-subjects variable being the selection method: Unimodal (dwell time) vs Multimodal (pointing then pressing the controller's trigger). This condition was counter-balanced with a Latin square.

We measured the effect on the following dependent variables:

- **Entry time**, measured from the moment the user pointed at the first object, until the moment the last object was selected. Only correct entries were included in the analysis of entry time to avoid skewing the results towards faster speeds due to including potentially aborted attempts, in line with related work [16, 48].
- **Error rate**, measured as the percentage of incorrectly entered passwords.

### 5.2 Procedure, Apparatus and Particiapnts

First, participants were explained the study and asked to fill a consent form and a demographics questionnaire. They were then asked to stand in the middle of the room and put on the HTC Vive. Participants who used the gaze pointing method underwent a calibration procedure at this point. Participants then had a training session in which they entered a 4-symbol password per condition to become acquainted to the system. Trial attempts were excluded from analysis. Participants examined a virtual board that showed the password they need to enter, and then entered the password three times successfully using the respective selection method. After the trial runs, participants entered three passwords three times each for each condition. The passwords were randomly generated prior to the study, and were displayed on the virtual board for participants to see before they started the authentication procedure. To ensure comparable results, the same set of passwords was used across all participants.

After each condition, we collected subjective feedback using likert-scale questions and a semi-structured interview. To evaluate memorability, participants were asked to define their own password to memorize at the end of the study, and were contacted one week later to fill in a follow-up questionnaire. The experiment complied with the university's ethics regulations.

We recruited 48 participants (14 females) aged between 18 and 57 years ($Mean = 25.23$, $SD = 7.35$) through mailing lists and social networks: 24 participants (4 females) used the gaze-based pointing method, and the other 24 (10 females) used the head pose pointing method. They were compensated with an online shopping voucher, and all had normal or corrected to normal vision.

### 5.3 Results

In total, we analyzed 864 password entries (3 passwords $\times$ 3 repetitions $\times$ 2 pointing methods $\times$ 2 selection methods $\times$ 24 participants per pointing method).

**Entry time** Figure 5A shows the overall entry time across all three entered passwords for each selection method in seconds. The figure emphasizes that both MultimodalGaze ($Mean = 5.94\,s$, $SD = 2.85\,s$) and MultimodalHead ($Mean = 5.51\,s$, $SD = 2.10\,s$) are faster than UnimodalGaze ($Mean = 9.15\,s$, $SD = 3.75\,s$) and UnimodalHead ($Mean = 9.56\,s$, $SD = 6.51\,s$). Indeed, a repeated

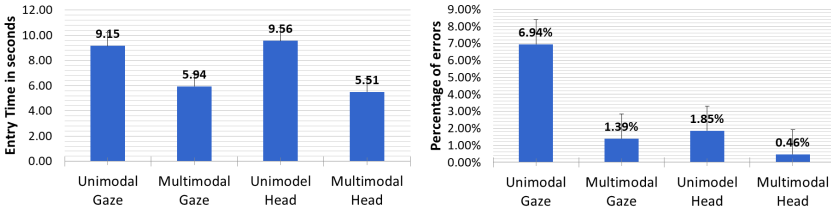


**Fig. 5.** In GazeRoomLock, multimodal selection is significantly faster than unimodal. We could not confirm that the choice of head or gaze for pointing impacts selection time ($p > 0.05$). Overall, participants performed very few errors using GazeRoomLock; less than 2% of entries using MultimodalGaze, UnimodalHead, and MultimodalGaze are erroneous. However, UnimodalGaze is more error-prone (6.94% error rate).

measures ANOVA revealed a significant effect of *selection method* on entry time ($F_{1,46} = 35.61$, $p < 0.001$). Post-hoc analysis using Bonferroni corrected t-tests indicated that unimodal selection ($Mean = 9.36\,s$, $SD = 0.56\,s$) is significantly slower than multimodal selection ($Mean = 5.72\,s$, $SD = 0.30\,s$). No significant effects of *pointing method*, or *repetitions* on entry time were found ($p > 0.05$). No interaction was found between pointing and selection method ($p > 0.05$).

This means multimodal selection is significantly faster than unimodal selection. The mean entry time using gaze is slightly faster than that of head-pose, but we have no evidence that one pointing method is significantly faster than the other.

**Error per Entry** Out of all 864 entries, only 23 were incorrect. 15 incorrect entries occurred in UnimodalGaze, 4 in UnimodalHead, 3 in MultimodalGaze and 1 in MultimodalHead. Figure 5B shows the average error rate for each *pointing method* (UnimodalGaze: 6.94%, UnimodalHead: 1.85%, MultimodalGaze: 1.39% and MultimodalHead: 0.46%). Note that most errors occurred when using UnimodalGaze. The overall error rate is 2.66%, which is lower than that in prior work on multimodal authentication [27] and authentication in IVR [16, 18].

Errors occurred either due to entering the password in a wrong order or switching the objects. The higher error rate in UnimodalGaze is attributed to the shaking of the headset, which leads to imprecise eye tracking.

A repeated measures ANOVA was used to analyze the errors. *Pointing method* was found to have a significant effect on error rate ($F_{1,46} = 8.76$, $p < 0.05$). Post-hoc analysis using Bonferroni corrected t-tests showed that Gaze ($Mean = 0.38$, $SD = 0.07$) is significantly more error prone than Head-pose ($Mean = 0.10$, $SD = 0.07$), ($p < 0.05$). Additionally, the *selection method* had a significant effect on error rate ($F_{1,46} = 8.7$, $p < 0.05$). Post-hoc analysis with Bonferroni corrected t-tests showed significant differences between the modalities, with unimodal selection ($Mean = 0.40$, $SD = 0.01$) being significantly more error prone than multimodal selection ($Mean = 0.08$, $SD = 004$), ($p < 0.05$). No significant effect for *repetitions* was found on error rate ($p > 0.05$). No interaction was found between pointing and selecting method ($p > 0.05$).

In summary, head-pose selections are significantly less error prone than gaze-based selections, and multimodal selection using the controller is significantly less error prone than unimodal selection using dwell time.

**Subjective Feedback** Participants preferred multimodal over unimodal approaches (83.3% preferred MultimodalGaze over UnimodalGaze, and 66.7% preferred MultimodalHead over UnimodalHead) due to their faster selection speed and the better control over input. Some unintentionally pointed away from the target before selecting it. Participants suggested using progress bars to show how many objects were selected so far as part of the password.

We used Generalized Estimating Equations (GEE) to analyze the Likert ratings [9]. This allowed us to take into account that we have ordinal dependent variables and both a within-subject factor (selection) and a between-subject factor (pointing). We found significant influences for the two Likert items *comfortable* and *not error-prone*: For the *comfortable* item, *selection* was a significant predictor; the odds of giving a higher *comfortable* rating with multimodal selection were 3.1 times the odds of unimodal selection ($p < 0.05$). Similarly, for the *not error-prone* item, *selection* was a significant predictor; the odds of a higher rating with multimodal selection were 3.4 times the odds of unimodal selection ($p < 0.01$). Moreover, *pointing* was also a significant predictor: The odds of a higher Likert rating on this item with head were 3.9 times the odds with gaze ($p < 0.01$).

**Memorability** As mentioned in section 5.2, participants were requested to define their password with a length of four objects at the end of the usability study. Duplicate objects were allowed, albeit not consecutively. To investigate the memorability of GazeRoomLock, 42 of the participants completed a follow-up questionnaire one week afterwards ($Gaze = 21$, $Head = 21$).

First, participants recalled which objects they had selected and entered their self-defined password without any cues. If unsuccessful, they were given pictures of the virtual environment and objects and were allowed to provide a second and third guess. 59.5% (25 participants: 14 Gaze-group, and 11 Head-group) remembered the correct passwords on the first trial. Participants who did not recall their passwords remembered the objects but not their order. After all three trials 83.30% (35 participants: 17 Gaze-group and 18 Head-group) remembered their passwords. There is no evidence that one modality results in better memorability than the other ($p > 0.05$). Participants who recalled their password memorized them in a story-like structure: “First I eat the burger, then chips as a side dish, take a sip of cola, and then eat a cake for dessert”. Three participants ordered the objects alphabetically: “banana, burger, book, can”. One participant memorized the positions of the objects: “The object on the window, then the one on the table, near the board then again the object on the window”. One participant remembered the objects by their colors: ”green, blue, green, blue”.

## 6 Observation-resistance Evaluation

IVR users are unlikely to notice the presence of bystanders either due to the HMD blinding them from seeing the real world, or due to the high immersiveness of the IVR experience. Indeed, George et al. [18] found that bystanders are able to observe passwords entered by IVR users. This underlines the need to evaluate GazeRoomLock against observation attacks.

We conducted two security studies to evaluate the observation resistance of GazeRoomLock. We chose the following two threat models, since they realistically simulate possible observation attacks against our system:

**Threat Model 1 – Real World Observations:** Here, the adversary is in the real world and observes the user who is wearing the IVR HMD. This closely aligned with prior work which suggests family and friends to be casual attackers [21].

**Threat Model 2 – Offline Observations:** Here, the attacker has access to two resources that allow them to perform offline observations: 1) A **video** recording that shows the user's movements as they authenticate while wearing the HMD. In a real scenario, this video can be retrieved by recording the user, for example, with a smartphone camera. 2) Additionally, the attacker has access to the **virtual scene** in which the user normally authenticates. This simulates the case where the attacker has full knowledge and access to the virtual environment (e.g., an insider that can use the victim's HMD in their absence), and can exploit this together with the videos to refine the observations. The reason we consider both resources is that in a real scenario, an attacker who has a video recording of the authentication process can revisit the virtual scenes whenever the HMD is unattended.

In both cases, we assume the attacker is able to identify the start and end of the authentication procedure, as confirmed by prior work [17]. After observing it, the attacker then puts on the HMD and tries to enter the password in the user's absence (see Figure 4).

We evaluated each threat model in a separate study, but they both followed the same study design. The key idea in both studies was to invite participants to observe password entries and try to make 3 guesses. All participants were compensated with online shop vouchers. To encourage high performance, we raffled an additional voucher in each study such that participants who were successful the most in attacking passwords would have higher chances in winning. Both studies complied with the university's ethics regulations

### 6.1 Study Design

We followed a within-subjects experiment design with one independent variable: **Input method** with four conditions: UnimodalGaze, UnimodalHead, MultimodalGaze, and MultimodalHead.

Participants went through 4 blocks (one per input method), in each of which they observed 3 passwords entered using the current block's input method. The order of blocks was counter balanced using a Latin square.

### 6.2 Security Study 1: Resisting Real World Observations

The first security study focused on real world attacks. We invited 26 participants (7 females) aged between 19 and 33 ($Mean = 23.71$, $SD = 3.8$). Participants were invited individually. They first signed a consent form and were explained the study and the reward mechanism. Afterwards, the experimenter put on the HMD and started entering passwords. The participant (i.e., the attacker) first had a training round for each condition, where the experimenter was observed while entering a random password. Attackers then observed the experimenter as she entered 12 passwords (4 input methods $\times$ 3 unique passwords) and were provided with pen and paper to take notes. Note that all passwords were unique and predefined by the experimenter, i.e., no passwords were entered more than once in front of the same participant. The attacker was explicitly told when the experimenter started and stopped entering the password. After observing each password, the attacker had the chance to make up to three guesses of the password, and was asked to enter them in IVR, while wearing the HMD. The study was concluded with a semi-structured interview. To avoid priming the participants, we did not reveal their performance until the end.

We analyzed 312 attacks (26 participants $\times$ 4 input methods $\times$ 3 unique passwords). Participant 3 and participant 10 were the only ones who failed in all their attacks. We believe they did not put enough effort as the rest of the participants who made at least one correct observation, and therefore we excluded their data to avoid skewing the results. Participants were successful in attacking only 20% of GazeRoomLock passwords. They were most successful in attacking UnimodalHead passwords, with a mean success rate of 33.33%, and least successful in attacking MultimodalGaze passwords with a mean success rate of 9.72%. Participants were equally successful in attacking UnimodalGaze and MultimodalHead passwords, with mean success rates of 18.06% for both of them. The results are illustrated in Figure 6.

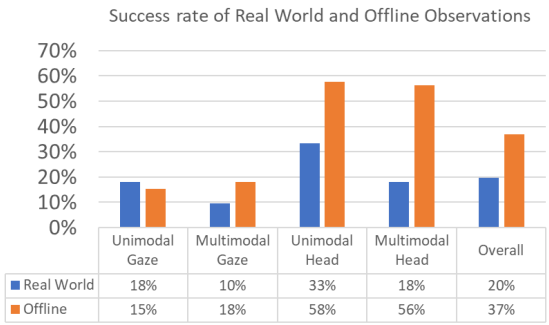


**Fig. 6.** The gaze-based methods are highly resilient to both types of observations. MultimodalHead is equally resilient to real world observations, but vulnerable to offline ones. UnimodalHead underperforms in both threat models.

### 6.3 Security Study 2: Resisting Offline Observations

The second security study focused on offline attacks. This study was conducted with a separate set of participants. We invited 26 different participants (7 females) aged between 18 and 31 ($Mean = 22.46$, $SD = 3.11$). After being introduced to the study and signing a consent form, participants were provided with a video showing the user entering a password in the real world (see Figure 4B), and a video of the virtual room without the user present. The user had full control over both videos (e.g., pause, rewind, etc.). For each attack, the participant was able to provide up to three guesses. Similar to the first security study, the attacker was given pen and paper.

We analyzed 312 attacks (26 participants $\times$ 4 input methods $\times$ 3 unique passwords). Participants were more successful in this threat model, with an overall success rate of 38.86% of GazeRoomLock passwords using offline attacks. Similar to the results of the previous study, participants were most successful in attacking UnimodalHead passwords. However the average success rate is much higher in case of offline attacks compared to real world attacks; offline attacks are successful on average 57.69% of the time in case of UnimodalHead, and almost similarly successful in case of MultimodalHead (56.41%). UnimodalGaze and MultimodalGaze were more resilient to offline attacks, with success rates as low as 15.38% and 17.95% respectively. Participants were most successful in offline attacks against passwords with targets within 90° (40.38%), less successful against passwords with targets within 180° (37.50%), and least successful against passwords with repetitions (32.96%).

**Limitations** We chose to investigate authentication using 4-symbols 3D passwords. This makes our results comparable to other works. We acknowledge that usability and security could differ based on the length of the password. For example, we expect that longer 3D passwords will take longer to enter, be more error prone, and be harder to observe. However, the relative results should not differ with different password lengths. This means that, for example, we expect longer passwords to be entered faster using MultimodalHead compared to UnimodalGaze.

## 7 Discussion

We evaluated the usability of GazeRoomLock in a user study, and we evaluated its observation resistance against real world and offline observations in two separate user studies.

### 7.1 Gaze-based Multimodal Authentication Improves Usability

We found that the multimodal selection approach significantly improves usability compared to a) the unimodal methods we evaluated and b) previous work in selecting 3D passwords.

**Faster Entry Time** As evidenced by the results, using an additional modality for authentication significantly reduces authentication time. In particular, users authenticated using MultimodalHead and MultimodalGaze in 5.51 s and 5.94 s respectively, but needed 9.56 s and 9.15 s when using UnimodalHead and UnimodalGaze. The reason behind the delay in the unimodal approaches is that users have to precisely point at the targets (i.e., dwell) for a period of time in order to select them. This is done in order to distinguish gaze that is intended for selection from the gaze behavior performed when scanning scenes – this is a challenge in gaze interfaces that is known as the Midas Touch [22]. While the dwell duration we used (800 ms) could theoretically allow selections in $4 \times 800$ ms = 3.6 s in addition to the time taken to find the objects, in practice our participants needed more time to maintain the gaze point in the object's collider. This is in part due to the inaccurate nature of gaze estimates [37]. While this can improve in the future by using better sensors and improved calibration algorithms, estimating the exact point of gaze is nearly impossible due to the physiology of the eye – the eye can fixate at 2° at a time, and the user can move attention in this area without any additional eye movements [37]. On the other hand, using an additional modality allows users to make selections as soon as they are pointing at the target, which allows for faster selections in VR.

Similarly, using the controllers only to enter passwords consisting of four 3D objects required 8.58 s and 14.32 s in prior work [16]. This means that entry time using multimodal approaches outperforms previous work on authentication mechanisms by 30.8% to 61.5%. We attribute the improved entry time to the fact that humans move their heads and eyes naturally faster than they point with their arms (see discussion on gaze vs pointing in [41]). Notably, 2D entry schemes, such as patterns, provide a better entry time (3 s) [48, 18]. However, this is due to the combined novelty of the device, input technique and password scheme. We are planning to investigate this further in a long-term study.

**Lower Error Rate** The multimodal selection methods also outperform the unimodal ones in terms of error rate (0.46% and 1.39% vs 1.85% and 6.94%). We expect that the difficulty of maintaining gaze at the target is the main reason behind the relatively higher error rate in the gaze-based conditions. Kytö et al. [33], who found similar trends in their results, suggest that their results for gaze-based interaction for AR may be transferred to VR, which we can confirm in the context of 3D password entry. These combined results suggest that authentication concepts are transferable between varying degrees of virtual reality and real world interaction.

Additionally, the multimodal approaches outperform previous work where controllers were used to enter 3D passwords at an error rate of 2.78% [16]. Methods that allow pointing with a controller require eye-hand coordination [8, 40], which can be a cognitively demanding process [40], and could thereby be the reason behind making more errors.

### 7.2 Resilience to Observations

Multimodal authentication offers higher observation resistance because it requires attackers to observe a) the user's pointing, and b) the user's selection. Splitting the attacker's attention is a strategy that overwhelms attackers. Previous work on authentication on mobile devices and public displays employed this strategy to complicate real time observations [47, 11]. However this strategy is not effective against video observations [30, 47].

Similar to previous work, results from our studies show that MultimodalHead is highly resilient to real world observations (18%) but vulnerable to offline observations (56%). On the other hand, MultimodalGaze is resilient to both real world (10%) and offline video observations (18%). In addition to splitting the attacker's attention, another reason the multimodal methods are resilient to observations is that they are fast; giving attackers less time to observe inputs.

UnimodalGaze is also resilient to both types of observations (online 15% vs real world 18%), but it is not recommended due to its low usability. Gaze-based methods are less vulnerable to video observations because of the subtle nature of eye movements, which are further obscured by the HMDs. On the contrary, head movements are more visible (58%).

### 7.3 Eye Tracker Calibration

In order to estimate a precise gaze point, users need to calibrate the eye tracker. Calibration is a procedure to map the user's eye movements, which are unique for every user, to points in the environment [37]. Calibration is perceived to be a tedious and time-consuming task [45]. The negative impact of calibration on HMD users was suggested to be negligible [29] because users do it only once, as opposed to desktop settings where it needs to be repeated whenever the setup is changed (e.g., change in user's, eye tracker's, or display's position). Previous work proposed alternative gaze input methods that do not require calibration (e.g., Pursuits [45] and gaze gestures [12]). While promising for calibration-free authentication, these methods typically require longer input times [28, 30].

### 7.4 Story-like 3D passwords

Regardless of input method, the majority of participants used story-like structures to memorize their passwords (e.g., *I ate the chips, and then had a cake for dessert*). Users link objects to tell "stories", facilitating memorability. This is similar to the concept of PassPhrases, where users attach multiple sentences [25]. This strategy has implications on usability and security. In terms of usability, it helps participants remember the passwords as shown in the study. On the other hand, obvious stories could make passwords predictable and result in easier guessing attacks. For example, knowing that cake is the only object described as a dessert, an attacker could predict that the last entry is the cake. A possible approach to counter this is to choose a diverse mix of objects and object categories to increase the possible options for potential stories.

### 7.5 Enabling a Choice of Input Modalities

Multimodal interaction methods resulted in better entry times and were found to be more "comfortable" to use in the presence of others. This confirms prior work by Alallah et al. [2] who investigated social acceptability of interaction methods for IVR. Their results show that users prefer controller based input methods, such as a touchpad, rather than mid-air gestures. Similarly, prior work on authentication for public displays, where real world bystanders are also a prominent threat, showed that users find mid-air gestures embarrassing to perform in public [7, 30]. These results suggest the need to provide the user with input modalities that adapt automatically depending on the context of the interaction or that enable the user to choose an appropriate one; for example in the presence of others vs. alone at home. Our results show that MultimodalGaze is not only more secure against offline attacks compared to the other methods, but also almost as highly usable as MultimodalHead. An HMD user in a public context may prefer using MultimodalGaze or even UnimodalGaze due to the inconspicuousness of gaze in public, rather than using more visible input methods, such as MultimodalHead. Outside a lab setting it is likely more challenging to identify the beginning and end of input when gaze is used, thereby granting even higher observation resistance. Therefore, we recommend that users are given the option to choose which multimodal method to authenticate with.

### 7.6 Future Work

In future work, we aim to explore calibration-free techniques for authentication in IVR. We also plan to investigate gaze input for established 2D authentication mechanisms, such as PIN and pattern in IVR. Furthermore, we plan to conduct a field study of GazeRoomLock (e.g., to allow HTC Vive users to log into their accounts). Field studies are becoming increasingly feasible now that many HMDs come with integrated eye trackers, such as the HTC Vive pro. Through the field study, we plan to better understand the effect of object location on password creation and how this impacts usability and security. We also plan to investigate strength meters for 3D passwords, and explore modalities beyond the controller's trigger (e.g., feet tapping). Furthermore, there are many ways the tactile input in our implementation can be improved. For example, Yang et al. [46] proposed subtle haptic techniques that can be very difficult to notice by bystanders; these methods can be used to replace the trigger in the multimodal approaches.

## 8 Conclusion

In this work we investigated multiple methods for pointing and selecting 3D passwords in immersive virtual reality: UnimodalGaze, MultimodalGaze, UnimodalHead, and MultimodalHead. Through three user studies (N=48, N=26, N=26), we investigated the usability of the approaches in terms of entry time, error rate

and memorability, and the security in terms of resistance to real world observations and offline observations. We found that the multimodal approaches are significantly faster and significantly less error-prone than the unimodal ones, while memorability does not change significantly depending on the pointing method. MultimodalHead is highly resilient to real world observations, but not offline observations, while MultimodalGaze is highly secure against both. We discussed how multimodal authentication can significantly improve usability and observation resistance over prior work. Rather than exclusively focusing on new authentication mechanisms, our work highlights that there is potential to improve usability and security of existing ones by adapting alternative input modalities for IVR.

## 9 Acknowledgements

The contributions from the authors Mohamed Khamis and Daniel Buschek were supported, in part, by the Royal Society of Edinburgh (Award number 65040), and the Bavarian State Ministry of Science and the Arts in the framework of the Centre Digitisation.Bavaria (ZD.B).

## References

1. 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: Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications. ETRA '19, ACM, New York, NY, USA (2019). https://doi.org/10.1145/3314111.3319837, http://doi.acm.org/10.1145/3314111.3319837
2. Alallah, F., Neshati, A., Sakamoto, Y., Hasan, K., Lank, E., Bunt, A., Irani, P.: Performer vs. observer: Whose comfort level should we consider when examining the social acceptability of input modalities for head-worn display? In: Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology. VRST '18, ACM, New York, NY, USA (2018). https://doi.org/10.1145/3281505.3281541, http://doi.acm.org.emedien.ub.uni-muenchen.de/10.1145/3281505.3281541
3. Alsulaiman, F., El Saddik, A.: Three-dimensional password for more secure authentication. Instrumentation and Measurement, IEEE Transactions on **57**, 1929 – 1938 (10 2008). https://doi.org/10.1109/TIM.2008.919905
4. Andrist, S., Gleicher, M., Mutlu, B.: Looking coordinated: Bidirectional gaze mechanisms for collaborative interaction with virtual characters. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. CHI '17, ACM, New York, NY, USA (2017). https://doi.org/10.1145/3025453.3026033, http://doi.acm.org/10.1145/3025453.3026033
5. Attree, E., Brooks, B., Rose, F., Andrews, T., Leadbetter, A., Clifford, B.: Memory processes and virtual environments: I can't remember what was there, but i can remember how i got there. implications for people with disabilities. In: ECDVRAT: 1st European Conference on Disability, Virtual Reality and Associated Technologies. Reading, UK. vol. 118 (1996)

6. Aviv, A.J., Gibson, K., Mossop, E., Blaze, M., Smith, J.M.: Smudge attacks on smartphone touch screens. In: Proceedings of the 4th USENIX Conference on Offensive Technologies. WOOT'10, USENIX Association, Berkeley, CA, USA (2010), http://dl.acm.org/citation.cfm?id=1925004.1925009
7. Brignull, H., Rogers, Y.: Enticing people to interact with large public displays in public spaces (2003)
8. Chan, L.W., Kao, H.S., Chen, M.Y., Lee, M.S., Hsu, J., Hung, Y.P.: Touching the void: Direct-touch interaction for intangible displays. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '10, ACM, New York, NY, USA (2010). https://doi.org/10.1145/1753326.1753725, http://doi.acm.org/10.1145/1753326.1753725
9. Clayton, D.: Repeated ordinal measurements: A generalised estimating equation approach (1992)
10. De Luca, A., Denzel, M., Hussmann, H.: Look into my eyes!: Can you guess my password? In: Proceedings of the 5th Symposium on Usable Privacy and Security. SOUPS '09, ACM, New York, NY, USA (2009). https://doi.org/10.1145/1572532.1572542, http://doi.acm.org/10.1145/1572532.1572542
11. De Luca, A., Harbach, M., von Zezschwitz, E., Maurer, M.E., Slawik, B.E., Hussmann, H., Smith, M.: Now you see me, now you don't: Protecting smartphone authentication from shoulder surfers. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '14, ACM, New York, NY, USA (2014). https://doi.org/10.1145/2556288.2557097, http://doi.acm.org/10.1145/2556288.2557097
12. Drewes, H., Schmidt, A.: Interacting with the computer using gaze gestures. In: Human-Computer Interaction – INTERACT 2007. pp. 475–488. Springer Berlin Heidelberg, Berlin, Heidelberg (2007)
13. Esteves, A., Velloso, E., Bulling, A., Gellersen, H.: Orbits: Gaze interaction for smart watches using smooth pursuit eye movements. In: Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology. UIST '15, ACM, New York, NY, USA (2015). https://doi.org/10.1145/2807442.2807499, http://doi.acm.org/10.1145/2807442.2807499
14. Esteves, A., Verweij, D., Suraiya, L., Islam, R., Lee, Y., Oakley, I.: Smoothmoves: Smooth pursuits head movements for augmented reality. In: Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology. UIST '17, ACM, New York, NY, USA (2017). https://doi.org/10.1145/3126594.3126616, http://doi.acm.org/10.1145/3126594.3126616
15. Forget, A., Chiasson, S., Biddle, R.: Shoulder-surfing resistance with eye-gaze entry in cued-recall graphical passwords. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '10, ACM, New York, NY, USA (2010). https://doi.org/10.1145/1753326.1753491, http://doi.acm.org/10.1145/1753326.1753491
16. George, C., Buschek, D., Khamis, M., Hussmann, H.: Investigating the third dimension for authentication in immersive virtual reality and in the real world. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (2019)
17. George, C., Janssen, P., Heuss, D., Alt, F.: Should i interrupt or not?: Understanding interruptions in head-mounted display settings. In: Proceedings of the 2019 on Designing Interactive Systems Conference. DIS '19, ACM, New York, NY, USA (2019). https://doi.org/10.1145/3322276.3322363, http://doi.acm.org/10.1145/3322276.3322363

18. 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: Proceedings of the Network and Distributed System Security Symposium (NDSS 2017). USEC '17, Internet Society (2017). https://doi.org/10.14722/usec.2017.23028, http://dx.doi.org/10.14722/usec.2017.23028
19. Gugenheimer, J., Mai, C., Mcgill, M., Williamson, J.R., Steinicke, F., Perlin, K.: Challenges using head-mounted displays in shared and social spaces. In: Proceedings of the 37th Annual ACM Conference on Human Factors in Computing Systems. CHI EA '19, ACM, New York, NY, USA (2019)
20. Gurary, J., Zhu, Y., Fu, H.: Leveraging 3d benefits for authentication. International Journal of Communications, Network and System Sciences **10**, 324–338 (01 2017). https://doi.org/10.4236/ijcns.2017.108B035
21. Harbach, M., von Zezschwitz, E., Fichtner, A., Luca, A.D., Smith, M.: It's a hard lock life: A field study of smartphone (un)locking behavior and risk perception. In: Symposium On Usable Privacy and Security (SOUPS 2014). pp. 213–230. USENIX Association, Menlo Park, CA (Jul 2014), https://www.usenix.org/conference/soups2014/proceedings/presentation/harbach
22. Jacob, R.J.K.: The use of eye movements in human-computer interaction techniques: What you look at is what you get. ACM Trans. Inf. Syst. **9**(2) (Apr 1991). https://doi.org/10.1145/123078.128728, http://doi.acm.org/10.1145/123078.128728
23. John, B., Koppal, S., Jain, E.: Eyeveil: Degrading iris authentication in eye tracking headsets. In: Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications. ETRA '19, ACM, New York, NY, USA (2019). https://doi.org/10.1145/3314111.3319816, http://doi.acm.org/10.1145/3314111.3319816
24. 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: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. p. 1–21. CHI '20, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3313831.3376840, https://doi.org/10.1145/3313831.3376840
25. Keith, M., Shao, B., Steinbart, P.: A behavioral analysis of passphrase design and effectiveness. Journal of the Association for Information Systems **10**(2) (2009), https://aisel.aisnet.org/jais/vol10/iss2/2
26. Ken Pfeuffer, Matthias J Geiger, S.P.L.M.D.B.F.A.: Behavioural biometrics in vr: Identifying people from body motion and relations in virtual reality. In: Proceedings of the 37th Annual ACM Conference on Human Factors in Computing Systems. CHI '19, ACM, New York, NY, USA (2019). https://doi.org/10.1145/3290605.3300340, https://doi.org/10.1145/3290605.3300340
27. Khamis, M., Alt, F., Hassib, M., von Zezschwitz, E., Hasholzner, R., Bulling, A.: Gazetouchpass: Multimodal authentication using gaze and touch on mobile devices. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. CHI EA '16, ACM, New York, NY, USA (2016). https://doi.org/10.1145/2851581.2892314, http://doi.acm.org/10.1145/2851581.2892314
28. Khamis, M., Hassib, M., Zezschwitz, E.v., Bulling, A., Alt, F.: Gazetouchpin: Protecting sensitive data on mobile devices using secure

multimodal authentication. In: Proceedings of the 19th ACM International Conference on Multimodal Interaction. ICMI 2017, ACM, New York, NY, USA (2017). https://doi.org/10.1145/3136755.3136809, http://doi.acm.org/10.1145/3136755.3136809

29. Khamis, M., Oechsner, C., Alt, F., Bulling, A.: Vrpursuits: Interaction in virtual reality using smooth pursuit eye movements. In: Proceedings of the 2018 International Conference on Advanced Visual Interfaces. AVI '18, ACM, New York, NY, USA (2018)
30. Khamis, M., Trotter, L., Mäkelä, V., von Zezschwitz, E., Le, J., Bulling, A., Alt, F.: Cueauth: Comparing touch, mid-air gestures, and gaze for cue-based authentication on situated displays. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. **2**(4) (Dec 2018). https://doi.org/10.1145/3287052, https://doi.org/10.1145/3287052
31. Kinnunen, T., Sedlak, F., Bednarik, R.: Towards task-independent person authentication using eye movement signals. In: Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications. ETRA '10, ACM, New York, NY, USA (2010). https://doi.org/10.1145/1743666.1743712, http://doi.acm.org/10.1145/1743666.1743712
32. Kumar, M., Garfinkel, T., Boneh, D., Winograd, T.: Reducing shoulder-surfing by using gaze-based password entry. In: Proceedings of the 3rd Symposium on Usable Privacy and Security. SOUPS '07, ACM, New York, NY, USA (2007). https://doi.org/10.1145/1280680.1280683, http://doi.acm.org/10.1145/1280680.1280683
33. Kytö, M., Ens, B., Piumsomboon, T., Lee, G.A., Billinghurst, M.: Pinpointing: Precise head- and eye-based target selection for augmented reality. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18, ACM, New York, NY, USA (2018). https://doi.org/10.1145/3173574.3173655, http://doi.acm.org.emedien.ub.uni-muenchen.de/10.1145/3173574.3173655
34. Labs, P.: Htc vive eye tracking add on. https://pupil-labs.com/blog/2016-08/htc-vive-eye-tracking-add-on/ (2016), accessed 01 April 2019
35. Mai, C., Khamis, M.: Public hmds: Modeling and understanding user behavior around public head-mounted displays. In: Proceedings of the 7th ACM International Symposium on Pervasive Displays. PerDis '18, ACM, New York, NY, USA (2018). https://doi.org/10.1145/3205873.3205879, http://doi.acm.org/10.1145/3205873.3205879
36. Majaranta, P., Aula, A., Räihä, K.J.: Effects of feedback on eye typing with a short dwell time. In: Proceedings of the 2004 Symposium on Eye Tracking Research & Applications. ETRA '04, ACM, New York, NY, USA (2004). https://doi.org/10.1145/968363.968390, http://doi.acm.org/10.1145/968363.968390
37. Majaranta, P., Bulling, A.: Eye tracking and eye-based human–computer interaction. In: Advances in physiological computing, pp. 39–65. Springer (2014)
38. Majaranta, P., Räihä, K.J.: Twenty years of eye typing: Systems and design issues. In: Proceedings of the 2002 Symposium on Eye Tracking Research & Applications. ETRA '02, ACM, New York, NY, USA (2002). https://doi.org/10.1145/507072.507076, http://doi.acm.org/10.1145/507072.507076
39. Rubin, P.: Review: Oculus go (January 2018, Lastchecked: 2019-02-07), https://www.wired.com/review/oculus-go/

40. Scrocca, M., Ruaro, N., Occhiuto, D., Garzotto, F.: Jazzy: Leveraging virtual reality layers for hand-eye coordination in users with amblyopia. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. CHI EA '18, ACM, New York, NY, USA (2018). https://doi.org/10.1145/3170427.3188618, http://doi.acm.org/10.1145/3170427.3188618
41. Sibert, L.E., Jacob, R.J.K.: Evaluation of eye gaze interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '00, ACM, New York, NY, USA (2000). https://doi.org/10.1145/332040.332445, http://doi.acm.org/10.1145/332040.332445
42. Sluganovic, I., Roeschlin, M., Rasmussen, K.B., Martinovic, I.: Using reflexive eye movements for fast challenge-response authentication. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. CCS '16, ACM, New York, NY, USA (2016). https://doi.org/10.1145/2976749.2978311, http://doi.acm.org/10.1145/2976749.2978311
43. Song, C., Wang, A., Ren, K., Xu, W.: "eyeveri: A secure and usable approach for smartphone user authentication". In: IEEE International Conference on Computer Communication (INFOCOM'16). pp. 1 – 9. San Francisco, California (April 2016)
44. Summers, N.: Microsoft's mixed reality hololens 2 headset is official (February 2019, Lastchecked: 2019-28-02), https://www.engadget.com/2019/02/24/microsoft-hololens-2-announced/
45. Vidal, M., Bulling, A., Gellersen, H.: Pursuits: Spontaneous interaction with displays based on smooth pursuit eye movement and moving targets. In: Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing. UbiComp '13, ACM, New York, NY, USA (2013). https://doi.org/10.1145/2493432.2493477, http://doi.acm.org/10.1145/2493432.2493477
46. Yang, J.J., Horii, H., Thayer, A., Ballagas, R.: Vr grabbers: Ungrounded haptic retargeting for precision grabbing tools. In: Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology. UIST '18, ACM, New York, NY, USA (2018). https://doi.org/10.1145/3242587.3242643, http://doi.acm.org/10.1145/3242587.3242643
47. von Zezschwitz, E., De Luca, A., Brunkow, B., Hussmann, H.: Swipin: Fast and secure pin-entry on smartphones. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. CHI '15, ACM, New York, NY, USA (2015). https://doi.org/10.1145/2702123.2702212, http://doi.acm.org/10.1145/2702123.2702212
48. von Zezschwitz, E., Dunphy, P., De Luca, A.: 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. pp. 261–270. MobileHCI '13, ACM, New York, NY, USA (2013). https://doi.org/10.1145/2493190.2493231, http://doi.acm.org/10.1145/2493190.2493231