2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) | 979-8-3315-1647-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISMAR62088.2024.00049

# Design and Evaluation of Controller-based Raycasting Methods for Secure and Efficient Text Entry in Virtual Reality

Tingjie Wan[1][2] Liangyuting Zhang[1] Yunxin Xu[1] Katie Atkinson[2] Lingyun Yu[1]
Hai-Ning Liang[3]*

[1] School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, China
[2] Department of Computer Science, The University of Liverpool, Liverpool, UK
[3] Computational Media and Arts Thrust, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China

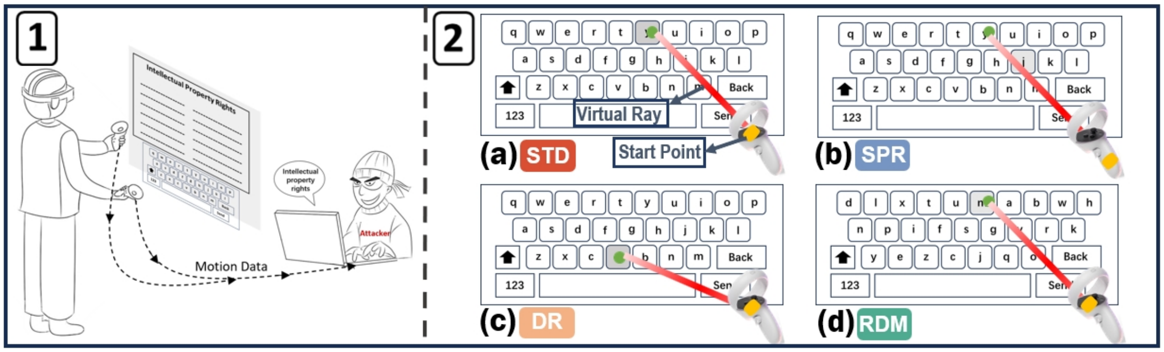

Figure 1: Figure 1⃞ is an example of input leakage, showing the attacker obtaining input text by remotely stealing the motion data of the controller. In Figure 2⃞, (a) The Standard Qwerty keyboard (STD, red). The two approaches for introducing variability in the virtual ray are (b) varying the start point (SPR, blue) and (c) altering the direction (DR, orange). (d) Random layout keyboard (RDM, green). In STD, the keyboard remains static. With SPR, the virtual ray's origin moves once following one key selection. DR entails a direction adjustment after each key selection. RDM presents a dynamically changing layout after completing a sentence.

## Abstract

With the exponential growth of digital information, ensuring text security, a fundamental component of information security, becomes increasingly paramount. While authentication remains a primary focus for data access control and protection, the rich sensor ecosystem and immersive experiences of virtual reality (VR) environments introduce new privacy risks, particularly with inconspicuous sensors like motion and location sensors. In this context, protecting the security of text entered by users poses a unique challenge. This paper explores the feasibility of enhancing text security by introducing variability in virtual input tools during typing processes. Specifically, we investigate the impact of introducing successive and random intermittent variations to the virtual ray (start point and direction) with controller-based raycasting techniques on text security and typing experience. The results demonstrate that introducing variability in virtual ray effectively protects regular text and passwords. Random intermittent introducing variability balances security and user experience for regular text. These findings provide insights into enhancing text security beyond authentication and defending against the potential risks in VR environments.

**Keywords:** Virtual reality; text entry; text security; keyboard layout; user study

## 1 Introduction

In today's digital era, the exponential growth of digital data underscores the critical importance of preserving the confidentiality of text. Confidential or private text often contains sensitive information, such as personal and financial records, business strategies, and classified government data. Unauthorized access to such text can result in severe repercussions, including identity theft, financial fraud, and breaches of national security. Therefore, it is imperative to implement robust privacy measures to mitigate these risks and maintain stakeholders' trust and confidence [14].

The discussion around information security often emphasizes authentication (such as password input [32, 41]) rather than text security [8]. Authentication serves as the initial defense for access control and data protection, ensuring that only authorized users can access systems or applications. However, even after authentication, devices with embedded sensors present significant security vulnerabilities. Previous studies have shown that user input data can be compromised by exploiting the device's sensors, such as accelerometers, gyroscopes, and motion sensors, to reveal the user's activities and location [29, 35, 38, 52, 53, 58].

VR devices encompass an array of sensors, ranging from overt ones like cameras and microphones to more inconspicuous ones like accelerometers and gyroscopes. While the privacy implications of widely recognized sensors such as cameras have garnered significant attention, some inconspicuous sensors often receive less scrutiny and protection [24]. In VR software development kits/APIs, accessing these inconspicuous sensors (e.g., motion, location, and button sensors) typically does not require secure permissions. This means that authentication alone in VR may not always effectively protect text from unauthorized access. This is particularly concerning given recent reports indicating that motion and location sensor data can be leveraged to estimate keystroke positions in 3D space, thereby

*Corresponding author (e-mail: hainingliang@hkust-gz.edu.cn)

This work is funded in part by the National Natural Science Foundation of China (#62272396).

979-8-3315-1647-5/24/$31.00 ©2024 IEEE
DOI 10.1109/ISMAR62088.2024.00049

revealing the geometric relationships between keystrokes [50,56].

The Qwerty keyboard stands as the most ubiquitous typing keyboard. Stealing user input is achieved by capturing the input actions and aligning them with the Qwerty keyboard layout [5]. Countermeasures against this keyboard alignment typically involve randomized key positions and auxiliary mappings [26,27,59,65]. While these methods effectively enhance text security, they introduce complexity during typing, resulting in unsatisfactory typing experiences. Schneider et al.'s study [41] found that the input rate for fully random layouts on physical Qwerty keyboards averaged 3.82 words per minute (WPM), contrasting with 21.0 WPM for non-random layouts. While slower input speeds may be tolerable in authentication scenarios involving password inputs, significant decreases in entry rates can notably impact user experience when inputting text especially regular text in VR. As regular text carries complete semantic information, multiple iterations of input may be necessary for various contexts that happen when composing text in messaging apps and editing documents. However, ensuring text security during text editing in VR poses an unexplored challenge. This necessitates investigation and development of approaches to maintain both speed and security in virtual environments.

In VR, virtual keyboards are a promising solution for natural and seamless interaction, as they could be integrated with dedicated controllers or advanced hand-tracking technologies, allowing convenient manipulation. Various input techniques employing virtual keyboards have been extensively evaluated in VR. In these techniques, the interaction with virtual keyboards does not involve direct manipulation using controllers or hands. Instead, virtual tools are often tethered to controllers (e.g., tethered to virtual rays [6,11,23,45,49,61,63]) or reconstructed as virtual hands (e.g., [12,17,37]) within the VR environment. This presents an opportunity to break the alignment between the keyboard layout and the user's input actions during typing in VR by introducing variations into the virtual input tools instead of the Qwerty keyboard. This is a cost-effective method to protect text security within VR environments, bypassing the necessity for extra protective measures like software or tools.

This study focuses on exploring the feasibility of introducing variability in virtual input tools during typing in VR to protect text (regular text and passwords). The controller-based raycasting technique represents a commonly employed text entry technique in VR. It has been applied to most commercial VR HMDs, which predominantly come with some type of handheld controller. Thus, we conducted our research based on the controller-based raycasting technique, a typical text entry technique associated with virtual selection tools, where the controller's front emits rays. We investigate the impact on security and typing experience of introducing variability in virtual rays (origin and direction). Our study is guided by two pertinent research questions:

- **RQ1** Can introducing variability in the virtual ray (start point and direction) bound on the controller, effectively ensure text security while meeting user expectations for their typing experience?
- **RQ2** How can typing security be enhanced, while meeting user expectations for their typing experience?

To address these RQs, our work began with a preliminary study (Section 4) to assess users' expectations regarding typing performance and security when entering regular text versus passwords and evaluated the feasibility of introducing variability in virtual rays (origin and direction) approaches for efficient and secure typing. These insights were instrumental in refining the design of secure and efficient text input. In the next user study (Section 6), we compared the effectiveness of introducing variability with successive frequency and random intermittent frequency to assess their ability to uphold text security while meeting user typing expectations.

Our findings revealed that introducing variability on virtual ray (start point and direction) can disrupt the mapping between input actions and keyboard layout, thus protecting the security of regular text and passwords. Introducing variation with random intermittent frequency (*ISPR* and *IDR*) exhibited better typing performance than those employing a successive introduction (*SPR* and *DR*). The main contributions of this paper are listed as follows: (1) We present a first systematic exploration of typing performance and security disparities between inputting regular text and passwords within VR environments; (2) We devise and evaluate four approaches of introducing variation in virtual ray, allowing for secure and efficient entry of text (two for password, two for regular text); and (3) We provide three recommendations for secure and efficient text entry design in VR.

## 2 Related work

### 2.1 VR Text Entry on Virtual Keyboard

In pursuit of portability and ecosystem uniformity, numerous studies have explored various text entry methods for typing on virtual keyboards in VR. These methods used different physical input devices and body parts, including handheld controllers [2,6,11,23,45,49,61–63], users' hands [12,17,37], head [30,48,57] eyes [40], and foot [51]. Typically, these techniques integrate physical input devices/body parts (controllers, hands, head, and eyes) with virtual tools to facilitate interaction with virtual keyboards in VR environments. Virtual rays are one of the most commonly used virtual selection tools in these text input techniques due to their linear nature, which makes operation in VR environments natural and intuitive. They can be emitted from either controllers or head/eye positions to indicate characters on a keyboard. Text input techniques utilizing virtual rays have shown relatively good performance, achieving up to 15.44 WPM when connected to controllers and up to 10.20 WPM when typing with head [45]. Another virtual linear selection tool is the virtual drumstick, which also performs well (21.01 wpm [6]). However, the virtual drumstick is currently only attempted to be connected to the controller due to its limited length. This is because the range of motion of hand-operated controllers exceeds that of the head and eyes, compensating for the limited length of the drumstick.

Another typical virtual selection tool is virtual hands, which replicate typing with both hands on a physical keyboard. Virtual hands require capturing real hands, typically using devices like cameras or gloves, and then reconstructing them into virtual representations for use in VR-based text input techniques. However, due to limitations in hand tracking accuracy and fatigue resulting from prolonged hand movements in the air, their current performance is less than ideal. For instance, in a study reported in Speicher et al.'s research [45], typing with two hands reached a maximum speed of only 9.77 WPM, which is lower than controller-based raycasting and drumstick input. Accurate tracking of hand movements often requires the use of external devices, such as the OptiTrack motion capture system utilized by Grubert et al. [17]. However, incorporating external devices for accurate hand tracking poses drawbacks, including added costs, portability issues, compatibility concerns, and security risks. Until recently, Dudley et al. [13] successfully utilized the built-in hand tracking of the Quest HMD to improve text entry performance to 20-30 WPM.

### 2.2 Risk of Input Leakage

There is significant potential for wearable devices with embedded sensors for monitoring and inferring human daily activities. However, these devices also pose serious security vulnerabilities, even when individuals are accessing key-based security systems [52]. This is primarily because inconspicuous sensors like accelerometers and gyroscopes often receive less scrutiny and protection compared to more prominent sensors such as cameras and microphones. For instance, studies have shown that accelerometers, gyroscopes, and motion sensors in smartwatches [29,53] and smartphones [35,38,58] can reveal the position of screen taps, thereby compromising user input data. Similarly, VR devices harbour similar security risks due to their reliance on various sensors. The study conducted by Wu et al. [56] has shown that it is possible to deduce user input content using motion sensor data collected from controllers on VR devices. Other researchers (e.g., [10, 16, 28, 47]) also confirmed that the hardware and software architecture of VR devices may be more susceptible to attacks or data theft, such as stealing input [28], breaking navigation [10].

Additionally, research by Shukla and Phoha [42] has demonstrated that attackers can exploit video recordings of hand movements to steal passwords entered on touchscreen smartphones. In the immersive environment of VR, users' awareness of their surroundings diminishes, thereby increasing the likelihood of unauthorized malicious activities like covert filming. Consequently, ensuring the privacy and security of user input in VR is crucial throughout the entire usage process, not just during the authentication stage.

## 3 Approaches of Introducing Variability in the Virtual Ray on Protection of Text

In controller-based raycasting text input techniques, the ray emitted from the controller serves as the selection tool for targeting keys. Adjusting the virtual ray's starting point and direction is a simple, low-cost, and easy-to-implement method for blurring user selections in VR. By dynamically altering the starting point or direction of the virtual ray, the angle of movement when using the controller for key selection remains changeable.

### 3.1 Introducing Variability in Start Point of the Ray (SPR)

Introducing variability in the start point of the ray (*SPR*) is an approach that involves adjusting the initial position of the ray associated with the controller, resulting in diverse rotation angles of the controller during object interactions (Figure 1b). The rotation angles are from the controller's built-in angle sensors. Following each confirmation of selection from the trigger button, the starting point of the ray is randomly shifted forward or backward. By modifying the starting point of the ray, this method introduces variability in the rotation angles necessary for object selection, even when the same key is chosen. When the starting point is adjusted, the cursor position on the keyboard remains unchanged. However, subsequent movements of the controller do not result in the same key presses as they would have if the SPR had not been modified.

### 3.2 Introducing Variability in Direction of the Ray (DR)

Introducing variability in the direction of the ray (*DR*) involves modifying the ray's direction with the controller (Figure 1c). By changing the ray's direction, this method enables variations in the rotation angles required for selecting objects, even when the same key or object is chosen. This adjustment is achieved by modifying the direction of the ray, thereby affecting the rotation angles needed for selecting objects. Even when starting with no initial rotation (i.e., a rotation angle of 0 degrees), the selected object might require some rotation during interaction due to the change in the ray's direction. The ray's direction was modified randomly once after each confirmation of selection from the trigger button.

## 4 Preliminary Study: Security and Performance Evaluation of Virtual Selection Tools Modification Approaches

The goal of this study is to conduct a comprehensive comparison and evaluation of two approaches introducing variations in virtual ray under different text types (password and regular text). To test the performance of our design under different text types, we included passwords from the RockYou password list [1] and sentences from a corpus of standardized English reading assessment [34]. To assess the efficacy and influence of the two approaches on both security and typing efficiency, we establish two baseline benchmarks: a standard keyboard layout (*STD*) and a randomly arranged keyboard (*RDM*). The *STD* layout represents the conventional and widely adopted configuration, serving as a reference point for text input performance (Figure 1a). On the other hand, the random layout keyboard (*RDM*) acts as a comparative measure, allowing us to evaluate the two approaches within a context where keys lack a specific arrangement (Figure 1e). These two baselines enable us to discern variations in security, text input performance, and user experience across different conditions.

The variation range of *SPR* is [-5, 5] units and the variation range of *DR* is [-16, 16] degrees in the horizontal direction and [-12, 12] degrees in the vertical direction. On the one hand, these values correspond to the wrist's maximum motion capacity, allowing about a 40-degree rotation within the horizontal plane [20]. On the other hand, we conducted pre-trials with the measuring method in Section 4.3.1 and the Identical Character Ratios (ICRs) of all examined passwords and the semantic similarity of all tested sentences remained below 30% when the variation range of *SPR* is [-5, 5] units and the variation range of *DR* is [-16, 16] degrees in the horizontal direction and [-12, 12] degrees in the vertical direction. Each time the amount of variation is at least greater than one key width required degree/unit.

### 4.1 Participants, Materials, and Apparatus

We recruited 16 participants (7 males and 9 females; aged 19 to 26, $M = 21.79, SD = 2.44$) from a local university. All participants reported familiarity with the Qwerty layout, as they regularly used laptops or desktop computers. Only three had no prior VR experience, seven had minimal VR exposure, fourteen were regular VR users, and ten had some experience typing in VR.

Our experiments utilized the Meta Quest 2 headset, boasting a rapid-switch LCD with 1832×1920 pixels per eye and a 90Hz refresh rate. This headset was paired with a Windows 10 PC housing an Intel i7-7700k CPU and an Nvidia GeForce GTX 1080 GPU. To establish our experimental setting and typing methods, we employed Unity3D (v2022.3.7f1c1) in conjunction with the Oculus XR Plugin (4.0.0) packages.

The virtual keyboard was positioned 10 meters ahead of the users, sized at 3.6 meters in width and 1.4 meters in height. Each character key remained uniform at 0.3 meters × 0.3 meters. These dimensions correspond to those utilized in earlier studies investigating text entry techniques.

In our pursuit of genuine and contextually relevant outcomes, we utilized the RockYou password list [1] and sentences from a corpus of standardized English reading assessment [34]. The RockYou password list, containing over 14 million plaintext passwords from the 2009 security breach, offers a realistic sample of commonly used passwords. It includes diverse combinations of upper and lowercase letters, numbers, and symbols, devoid of inherent linguistic meaning. The standardized English reading assessment [34] contains twenty 300-word passages and is written within the British National Corpus of 500 most frequently used words of English, which ensures the effectiveness of our assessment in facing regular language scenarios. Our focus is on the attacks that exploit large volumes of user input action data to decipher the entered content (see Section 2.2). By using the standardized English reading assessment, we can select longer sentences that contain richer contextual meanings and are typically closer to what users use daily.

### 4.2 Experiment Design and Procedure

This study employed a within-subjects design with Text Type and Approach as two independent variables. The experiment consisted of four sessions corresponding to the four approaches. In each session, participants needed to enter passwords and meaningful sentences with one approach (*STD*, *RDM*, *SPR*, or *DR*). A Latin Square design counterbalanced the order of APPROACH and TEXT TYPE. In each session, participants needed to complete 10 sentences for each approach. Each session had 10 sentences with no duplicates, five randomly selected from the RockYou password list [1] and five randomly from the standardized English reading assessment [34]. Before each session, participants were given 1 sentence and 1 password for training. The chosen password consists of 6 to 8 characters, comprising two different character types (letters, symbols, or numbers). This type of password is common in the RockYou password dataset. Sentences from the standardized English reading assessment require more than 100 keystrokes. Regular text inputs typically encompass more than a few characters. For instance, In tasks such as composing emails or engaging in social media interactions, users often input multiple sentences or paragraphs of text to convey their messages effectively. Participants need to input three different full-letter sentences before entering the target sentence. This helps simulate sufficient keystroke information that can be gained from realistic regular text input scenarios. This led to a total of 640 trials used for further analyses (= 16 participants × 4 approaches × (5 passwords + 5 sentences)).

Before starting, participants completed a consent form and a demographics questionnaire. Next, they were familiarized with the VR device, task objectives, and controls, emphasizing both speed and accuracy in typing. The experiment included four sessions, each with a specific approach and two types of text. Post-task questionnaires followed each session. After completing all sessions, participants participated in semi-structured interviews for feedback and suggestions. A five-minute break was provided between sessions, with extra time upon request. The entire experiment lasted about 90 minutes.

## 4.3 Evaluation Metrics

### 4.3.1 Performance of Protecting Input Content

We assessed the efficacy of the four approaches in ensuring security by analyzing the sensor data gathered in the study. We recorded motion data each time the trigger button on the controller was pressed for key selection, capturing pitch and roll angles, position, and the Z-axis distance between the keyboard and the controller. Employing the algorithms outlined in Wu et al.'s study [56], we replicated the 3D cursor estimation and K-means clustering techniques to estimate the 3D cursor's location and adjust the 2D keyboard layout based on the recorded data. Additionally, we replicated Wu et al.'s methods for password and paragraph predictions [56]. For password prediction, we utilized a tree-based backward typing trajectory for inference. Given that participants consistently concluded their input by pressing the "Send" key, we employed the reconstructed 2D keyboard alongside a tree-based backward inference algorithm. In the case of longer sentences, we clustered keystrokes using DBSCAN [19], aligned keyboards using LSE and marked keystrokes via KNN.

- **Identical Character Ratio** (ICR) calculates the proportion of identical characters at corresponding positions between two strings to the total number of characters. This measurement compares the absolute consistency between original and transcribed strings. It is particularly well-suited for passwords due to their stringent character order and consistency requirement.

- **Semantic Similarity** is defined as the similarity of two sentences in semantic space. The computational approach utilizes the cosine similarity [39] to assess the similarity between two sentences. This metric model captures sentence semantics by representing sentences as word vectors and then measuring the angle between these word vectors to gauge their similarity. The closer the cosine similarity is to 1, the more similar the two sentences are; the closer it is to 0, the less similar the two sentences are.

### 4.3.2 Performance of Text Entry

We measured the text entry performance of the four approaches under the raycasting technique using the objective data recorded during the experiments. Additionally, we gathered the subjective data from two questionnaires and an interview.

- **Entry Rate** was quantified in words per minute (WPM) [60]. This measure was derived by dividing the number of transcribed words by the time taken to complete the text transcription, measured in minutes. A word was defined as five keystrokes long, encompassing spaces.

- **Error Rate** [44] was evaluated employing char-level typing metrics, where the total error rate (TER) constitutes the combination of the not corrected error rate (NCER) and the corrected error rate (CER).

- **Workload** associated with the four approaches was evaluated using the NASA-TLX workload questionnaire [18]. This questionnaire comprises six subscales that gauge various facets of workload, encompassing mental demand, physical demand, temporal demand, frustration, effort, and performance. Participants provided ratings for each subscale on a scale from 0 to 100, with 5-point intervals. Reduced scores on this scale signify decreased workload and enhanced overall performance.

- **Usability** of each approach was measured using a System Usability Scale (SUS) questionnaire [9]. This questionnaire consists of 10 questions, each rated on a 5-point scale. The analysis involved using the weighted overall score, ranging from 0 to 100, where higher scores indicate enhanced usability.

- **Interview** was a semi-structured discussion regarding (1) participants' typing experience when using each approach; (2) their typing expectations when inputting passwords and sentences; and (3) any possible recommendations.

## 4.4 Results

We used SPSS 26 for data analysis. Shapiro-Wilk tests and Q-Q plots indicated that entry rate, SUS, and NASA-TLX data were normally distributed ($p > .05$), while ICR, semantic similarity, TER, and NCER were not normally distributed ($p < .05$). Thus, we applied Aligned Rank Transform [54] to non-normally distributed data before applying Repeated Measure (RM-) ANOVA tests. As there are six dimensions in NASA-TLX data, we used Multivariate ANOVA (MANOVA) to compare the differences. For one-dimensional normally distributed data, we applied RM-ANOVA tests.

### 4.4.1 Identical Character Ratio and Semantic Similarity

An RM-ANOVA analysis revealed a significant main effect of APPROACH on the ICR ($F_{3,45} = 147.316, p < .001, \eta_p^2 = .985$), and APPROACH × TEXT TYPE showed a significant difference ($F_{3,45} = 3.132, p = .035, \eta_p^2 = .428$) as illustrated in Figure 2a. The ICR of *STD* was significantly higher than *SPR*, *DR*, and *RDM* in both password and sentence (all $p < .001$). With *STD*, the ICR of password ($M = 69.28\%, SD = 9.04$) was significantly lower than that of sentence ($M = 79.79\%, SD = 7.81$) ($p = .005$).

An RM-ANOVA analysis indicated that APPROACH ($F_{3,45} = 127.724, p < .001, \eta_p^2 = .895$) had a significant effect on semantic similarity of the sentence, as illustrated in Figure 2b. Post-hoc pairwise comparisons indicated semantic similarity of *STD* was noticeably higher than *SPR*, *DR*, and *RDM* (all $p < .001$).

### 4.4.2 Entry Rate and Error Rate

An RM-ANOVA analysis indicated that both APPROACH ($F_{3,45} = 147.331, p < .001, \eta_p^2 = .905$) and TEXT TYPE ($F_{1,15} = 128.701, p < .001, \eta_p^2 = .889$) had significant main effects on entry

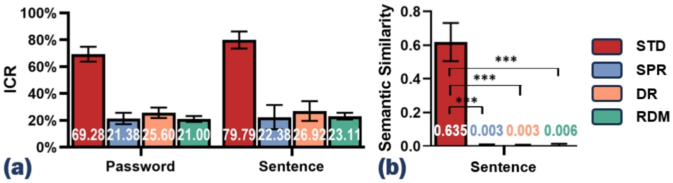

Figure 2: The means of (a) Identical Character Ratio and (b) Semantic Similarity of *STD*, *SPR*,*DR* and *RDM* in the preliminary study. Error bars represent 95% confidence intervals.

rate, and APPROACH × TEXT TYPE showed a significant difference ($F_{3,45} = 12.402, p < .001, \eta_p^2 = .528$), as illustrated in Figure 3a. Post-hoc tests showed the text entry rates for sentences surpassed those for passwords across the three Approaches (*STD*, *SPR* or *DR*) ($p < .001$). *STD* exhibited higher entry rates compared to the other approaches for both password and sentence typing ($p < .001$). *SPR* and *DR* showed significantly faster text entry than *RDM* when typing sentences and passwords (all $p < .001$). *DR* outperforms *SPR* when typing sentences ($p < .001$) and passwords ($p = .007$).

Figure 3b and c summarize the TER and NCER results, respectively. RM-ANOVAs analysis indicated that only TEXT TYPE ($F_{1,15} = 6.761, p = .020, \eta_p^2 = .678$) had a significant main effect on TER, as illustrated in Figure 3b. Post-hoc tests did not reveal any significant differences. RM-ANOVAs did not reveal any significant differences for NCER ($p > .05$).

### 4.4.3 Usability and Perceived Workload

An RM-ANOVA found a significant main effect of APPROACH in SUS scores ($F_{3,45} = 100.450, p < .001, \eta_p^2 = .8698$) (Figure 3d). Post-hoc pairwise comparisons revealed that the usability of *STD* ($M = 91.09, SD = 8.75$) was higher than it of *SPR* ($M = 64.88, SD = 9.23)(p = .002)$, *DR* ($M = 67.34, SD = 12.20)(p = .016)$ and *RDM* ($M = 39.53, SD = 14.27)(p < .001)$. *RDM* had lower usability than *SPR* ($p = .016$) and *DR* ($p = .002$).

Figure 3e shows the NASA-TLX scores for the four text entry approaches. MANOVAs revealed a significant difference in perceived workload ($F = 43.798, p < .001, Wilks'\Lambda = .037, \eta_p^2 = .963$). For each dimension of NASA-TLX, RM-ANOVAs showed significant effects in mental demand ($F_{3,45} = 69.176, p < .001, \eta_p^2 = .822$), physical demand ($F_{3,45} = 27.930, p < .001, \eta_p^2 = .651$), temporal demand ($F_{3,45} = 56.619, p < .001, \eta_p^2 = .791$), effort ($F_{3,45} = 25.482, p < .001, \eta_p^2 = .629$), performance ($F_{3,45} = 21.748, p < .001, \eta_p^2 = .592$), and frustration($F_{3,45} = 42.484, p < .001, \eta_p^2 = .793$). Post-hoc tests indicated that *STD* achieved the lowest workload, followed by *SPR* and *DR*, with *RDM* having the highest workload across all six dimensions.

### 4.4.4 Interview

Participants expressed that *STD* was the most user-friendly, as they could type on this virtual keyboard just like on a physical keyboard. A smooth typing experience means users did not noticeably feel the impact of password protection strategies used during typing, which would force them to slow down or pause the process. When using *SPR* and *DR*, participants reported experiencing noticeable pauses, which hindered their smooth input process. With *RDM*, they always had to expend considerable effort locating the target keys, which was the main reason for the inefficiency. Participants expressed that using *SPR* and *DR* for entering passwords is satisfactory. They also indicated that using *RDM* is acceptable, particularly if the frequency of password entry is low. Since passwords are typically short, even if the efficiency of inputting them is low, it would not consume too much time. However, when it comes to regular text input, especially for tasks involving heavy text such as writing reports, participants found it unacceptable if the efficiency is low and the user experience is uncomfortable. Therefore, participants stated that they were unwilling to use *RDM* for regular text input. While *SPR* and *DR* were considered somewhat more usable than *RDM*, they still found it frustrating, especially inputting heavy text.

## 4.5 Discussion (RQ1)

Our results demonstrate the effectiveness of introducing variability in virtual input tools in protecting text security within the controller-based raycasting text entry technique, since the ICRs for *SPR* and *DR* are significantly lower than that of *STD*, achieving a level of resistance comparable to *RDM* when entering a password and sentence (see Figure 2a). Additionally, the semantic similarity of sentence for *STD* was significantly higher than that of the three approaches (refer to Figure 2b).

The entry rate using *SPR* and *DR* remains lower than that of using the standard keyboard (*STD*) but higher than using the completely randomized design (*RDM*) (shown in Figure 3a). The retention of the standard virtual Qwerty keyboard ensures that users can retain the typing capabilities developed on this keyboard. In previous research (e.g., [2, 25, 30, 41, 48]), the Qwerty keyboard has often been highlighted, because the familiarity and widespread adoption of the Qwerty layout among users contribute significantly to its efficiency [21, 32]. However, introducing variations in the virtual ray's starting point/direction after each input led users to feel pauses or gaps during typing, resulting in an unsmooth input experience that affected their usability. Participant feedback supports this, with only *STD* exhibiting satisfactory usability, achieving an average SUS score of 91.09 (where 70 represents an acceptable level [3]). The SUS scores for *SPR* (64.88) and *DR* (67.34) fall below the acceptable level, while *RDM* score (39.53) was the lowest. Furthermore, participants reported a moderate level of cognitive involvement during typing tasks with *SPR* and *DR*, as indicated by workload ratings within the range of 30 (Figure 3e). In contrast, the workload for *RDM* was notably high, exceeding 40, which is considered a high workload. Error rates (TER and NCER) for both *SPR* and *DR* were similar to those observed with *STD* and *RDM*, suggesting The introduction of variability and keyboard layout had no significant impact on the error rates and user-initiated error correction. STD's TER is relatively higher compared to the other three conditions (Figure 3b). This is because the typing entry rate and error rate are interrelated. When typing speed increases and becomes more fluid, users are more prone to making errors. Conversely, when typing speed decreases, users have more time to accurately locate the target keys, leading to fewer mistakes. This is a common characteristic in typing tasks [64].

With *SPR*, the required controller movement for the same cursor distance is inconsistent due to changes in the starting point. The advantage of *SPR* is maintaining a consistent ray direction, ensuring the pointing position on the virtual keyboard remains unchanged despite variations in the starting point. Visual guidance promotes natural interaction [33, 36]. However, altering the starting point may weaken the connection between the ray and the controller, making the interaction less visually observable. When the ray's starting point is far from the controller, the ray becomes shorter and harder to discern, leading users to rely primarily on the visual cursor to perceive the controller's movement distance.

In *DR*, alteration of the ray's direction has a lesser impact on the sense of connection between the ray and the controller since the starting point remains unchanged. Users can continuously perceive the ray's direction, providing clues about the controller's movement trajectory. However, changes in the direction of the ray still result in the diversion of user attention, as changes in cursor position resulting from alterations in the direction of the virtual ray necessitate users to first determine the cursor's location before moving it, thereby

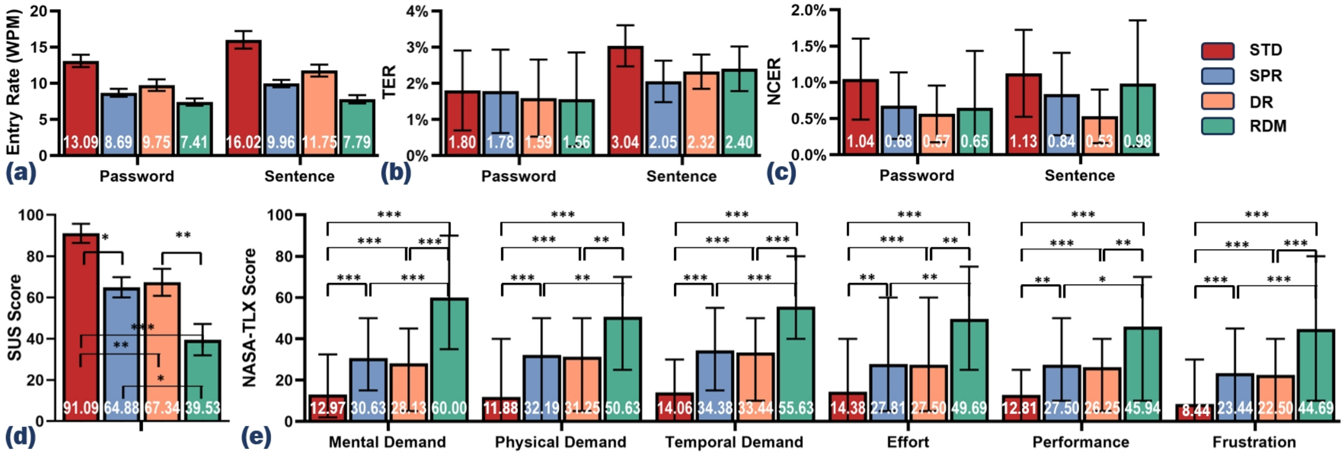

Figure 3: The means of (a) Entry Rate; (b) TER; (c) NCER; (d) SUS Scores (the higher, the better) and (e) NASA-TLX Scores (the lower, the better) of *STD, SPR,DR* and *RDM* in the preliminary study. Error bars represent 95% confidence intervals. In Figures d and e, ***, **, and * represent a .001, .01, and .05 significance level (Bonferroni-adjusted), respectively.

impacting the smoothness and efficiency of typing.

The entry rates for sentences are notably faster than for passwords when using the Qwerty keyboard (*STD*, *SPR* and *DR*) (see Figure 3a), consistent with previous findings [48]. Sentence inputs benefit from the contextual cues provided by the language, allowing users to leverage their familiarity with keyboard layouts and semantic understanding to achieve faster entry rates. Passwords typically lack meaningful context, prompting users to exercise greater caution during input. Commonly used words and letter combinations in the Qwerty keyboard often develop space memory, allowing users to input these words more quickly with less need for conscious thought and search [4]. Consequently, the entry rate of entering sentences (16.02 WPM) on a standard keyboard (*STD*) is significantly higher than that of entering passwords (13.09 WPM). *SPR* and *DR* introduce gaps on the standard keyboard, but the layout remains intact, enabling slightly higher sentence entry rates (*SPR*: 9.96 WPM, *DR*: 11.75 WPM) than passwords *SPR*: 8.69 WPM, *DR*: 9.75 WPM). In contrast, a random keyboard layout (*RDM*) provides no familiarity, depriving users of any benefits from layout familiarity and resulting in consistent but slower entry rates for both sentences (7.79 WPM) and passwords (7.41 WPM).

In summary, the impact on the text entry performance can be attributed to language context, keyboard familiarity, and smoothness of use. While *SPR* and *DR* offer some improvement over *RDM*, the frequent changes in the virtual ray's starting point and direction negatively impact the user's typing experience, leading to worse and less efficient typing experiences, especially for sentences, as participants expressed in interview 4.4.4.

### 4.6 Lessons

The following lessons (**L#**) were learned from this preliminary study.

**L1.** *SPR* and *DR* approaches are effective in protecting input text including passwords and sentences (**RQ.1**).

**L2.** Users can employ *SPR* and *DR* for limited text input (e.g., password), but for regular text, such as for online instant messaging chat or a document, it remains cumbersome and uncomfortable (**RQ.1**).

## 5 Refinement of *SPR* and *DR* for Regular Text

Both techniques have demonstrated effectiveness in preserving input content privacy. However, they also incur a partial decline in typing performance. This is primarily due to the successive variation introduction, necessitating constant adaptation by users to new rays after each key selection, as mentioned in Section 4.5.

In predicting regular texts such as sentences or paragraphs, a significant amount of keystroke data needs to be analyzed and processed to align keystroke behaviour with the keyboard layout, to establish the relationship between keystroke behaviour and text. This process involves categorizing keystroke behaviour and reconstructing the keyboard. Constructing a keyboard (rectangular structure) requires a minimum of 4 points that are not on the same row. Hence, reconstructing the keyboard necessitates at least 12 keystrokes, with each key requiring a minimum of three keystrokes for clustering. The 4 dots correspond to a keyboard containing 31 keys, for a total of 2430 total alignments (= 10 keys × 9 keys × 9 keys × 3 keys). The reverse inference method only applies to passwords, as passwords have limited lengths, and increasing length leads to exponential growth in computation.

Considering the features of predicting regular texts and the user experience of using two approaches to changing the virtual ray, we considered intermittent changes in the ray. By not immediately altering the ray after each keystroke, but instead setting a random interval frequency, the user's keystroke privacy may also be effectively protected, as long as this interval frequency is higher than the minimum data required to reconstruct the keyboard layout.

Additionally, by randomizing the interval frequency, we further enhance the challenge for potential attackers. We set the interval frequency within a random range of 4 to 12 occurrences. We modified the timing of introducing variations for the starting point/direction of the ray to occur randomly and intermittently, according to our above analysis. These schemes are named randomly and intermittently varying the start point of the ray (*ISPR*) and randomly and intermittently varying the direction of the ray (*IDR*).

## 6 User Study

The objective of this study is to assess the security, text entry performance, and user experience of *ISPR* and *IDR* (random and intermittent introducing variation) in comparison to *SPR* and *DR* (successive introducing variation).

### 6.1 Participants and Apparatus

A total of 24 participants were recruited (12 females, 12 males; aged 18 to 27, $M = 22.75, SD = 4.10$) from the same university. The apparatus was the same as that used in the last study. Participants were familiar with the Qwerty keyboard. Three had no prior VR experience, six had limited VR use without typing, and fifteen used VR several times a month, with twelve having some experience with VR typing systems. Ten participants also took part in the preliminary study. A three-week gap between the two studies was used to minimize learning effects.

### 6.2 Experiment Design and Procedure

This study utilized a within-subjects design, with APPROACH serving as the independent variable (*SPR*, *DR*, *ISPR* and *IDR*). The experiment consisted of four sessions corresponding to the four approaches. The order of the four sessions was counterbalanced using a Latin-Square approach. In each condition, participants transcribed 10 sentences sourced from the standardized English reading assessment [34]. Sentences were selected randomly without any duplicates. To simulate a more realistic long-text input scenario, participants are required to input three distinct full alphabet sentences before entering the target sentence. In total, 960 trials were included in the analysis (= 24 participants × 4 approaches × 10 sentences).

### 6.3 Results

We used SPSS 26 for data analysis. Shapiro-Wilk tests and Q-Q plots indicated that ICR, entry rate, TER, and SUS data were normally distributed ($p > .05$), while semantic similarity, NCER and NASA-TLX data were not normally distributed ($p < .05$). Thus, we applied Aligned Rank Transform [54] to NCER and NASA-TLX data before applying RM-ANOVA/MANOVA tests.

#### 6.3.1 Identical Character Ratio and Semantic Similarity

No significant effect of APPROACH was observed in both ICR and semantic similarity, as indicated by the results from RM-ANOVAs (both $p > .05$). Figure 4a and b summarize the ICR and semantic similarity respectively. The ICR for *DR* ($M = 19.93\%, SD = 2.48$), *IDR* ($M = 18.07\%, SD = 2.91$), *SPR* ($M = 20.19\%, SD = 4.85$), and *ISPR* ($M = 19.91\%, SD = 2.48$) all hovered around 20%, while the semantic similarity for the four approaches all were nearly 0%.

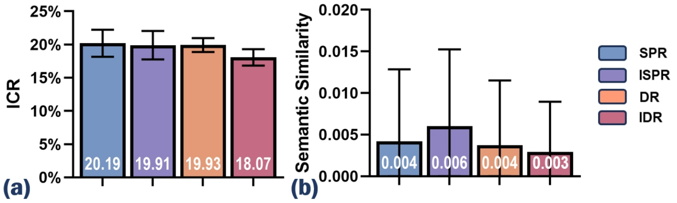

Figure 4: The means of (a) Identical Character Ratio and (b) Semantic Similarity The means of (a) Identical Character Ratio and (b) Semantic Similarity of *SPR*, *ISPR*,*DR* and *IDR* in the user study. Error bars represent 95% confidence intervals. ***, **, and * represent a .001, .01, and .05 significance level (Bonferroni-adjusted), respectively. The same marking scheme is used in Figure 5.

#### 6.3.2 Entry Rate and Error Rate

An RM-ANOVA analysis indicated that APPROACH ($F_{3,69} = 97.04, p < .001, \eta_p^2 = .815$) had a significant effect on entry rate, as illustrated in Figure 5a. Post-hoc pairwise comparisons indicated *IDR* ($M = 12.95, SD = 1.83$) and *ISPR* ($M = 12.45, SD = 1.53$) demonstrated the highest entry rates, followed by *DR* ($M = 10.75, SD = 1.55$). Meanwhile, *SPR* exhibited the slowest rate ($M = 8.52, SD = 1.02$). In all significantly different pairs, the significance level was less than .001.

The RM-ANOVAs demonstrated a statistically significant effect of APPROACH on TER ($F_{3,69} = 3.923, p = .012, \eta_p^2 = .161$), but no significant differences in NCER ($p > .05$) (Figure 5b and c). Post-hoc tests did not reveal any significant differences for TER.

#### 6.3.3 Usability and Perceived Workload

RM-ANOVA tests revealed significant differences in SUS score ($F_{3,69} = 29.31, p < .001, \eta_p^2 = .560$) among the four approaches (Figure 5d). Post-hoc tests revealed that *ISPR* ($M = 83.23, SD = 16.66$) and *IDR* ($M = 79.36, SD = 14.11$) generated the highest usability, succeeded by *DR* ($M = 67.71, SD = 12.00$). Conversely, *SPR* was the least usable ($M = 57.08, SD = 16.66$).

Figure 5e shows the NASA-TLX scores for the four approaches. MANOVAs revealed a significant difference in perceived workload ($F = 40.124, p < .001, Wilks'\Lambda = .070, \eta_p^2 = .930$). For each dimension of NASA-TLX, RM-ANOVAs showed significant effects in mental demand ($F_{3,69} = 11.571, p < .001, \eta_p^2 = .335$), physical demand ($F_{3,69} = 23.390, p < .001, \eta_p^2 = .504$), temporal demand ($F_{3,69} = 27.209, p < .001, \eta_p^2 = .542$), effort ($F_{3,69} = 23.167, p < .001, \eta_p^2 = .502$), performance ($F_{3,69} = 25.897, p < .001, \eta_p^2 = .530$), and frustration ($F_{3,69} = 26.271, p < .001, \eta_p^2 = .533$). Post-hoc tests indicated that the workload linked to *IDR* and *ISPR* was notably lower than that of *SPR* across all six dimensions ($p = .004$ in mental demand between *SPR* and *IDR*, $p < .001$ for the other pairs). The workload of *IDR* and *ISPR* was lower than that of *SPR* except for mental demand. Participants exerted more effort ($p = .037$), expressed greater dissatisfaction ($p = .008$), and experienced stronger frustration ($p = .011$) when using *IDR* compared to using *ISPR*. But the mental, physical, and temporal demands experienced with *ISPR* and *IDR* did not exhibit significant differences.

### 6.4 Discussion (RQ1)

The results of our study demonstrate the effectiveness of all four approaches in protecting text security, as evidenced by the low ICRs of about 20% and nearly zero semantic similarities (Figure 4). However, for text entry rates and user experience (usability and perceived workload), *ISPR* and *IDR* (i.e random and intermittent alteration) outperformed *SPR* and *DR* (i.e successive alteration), as illustrated in Figure 5a, d, and e.

Both random intermittent and successive variation introductions demonstrated comparable competence in the security of regular texts. Whether introducing variations successively or intermittently, the unpredictability of the mapping between input actions and characters makes it difficult for attackers to accurately infer characters or words, effectively protecting text security.

The entry rates of *ISPR* (12.45 WPM) and *IDR* (12.95 WPM) were close to the entry rate of *STD* observed in our previous study (16.02 WPM) and those reported for various raycasting techniques (e.g., 16.65 WPM [6] using MacKenzie's phrase set [31], 15.44 WPM [45], 17.4 WPM [63], 19.75 WPM [61] with the Enron mobile email dataset [46], and 15.94 WPM [49] with Brown Corpus [15]). Participants' subjective feedback also indicated that they experienced minimal disruption due to changes in the ray during typing, describing their typing experience as smooth. That was consistent with the SUS results that *ISPR* and *IDR* showed acceptable usability (mean score over 70 [3]), while the usability of *SPR* and *DR* was below the acceptable level. In Figure 4a, the ICRs of *ISPR* and *IDR* are lower than *SPR* and *DR* due to one instance in *SPR* and *DR* results where a prediction was close to 30%. This did not occur in *ISPR* and *IDR* results, leading to the results shown in this figure.

Participants reported that the changes in the starting point of the virtual ray in *ISPR* were very subtle, almost imperceptible, with only slight differences felt in hand movements. In contrast, *IDR* showed more noticeable changes in the direction of the virtual ray, making cursor movement within the field of view clearly visible. Also, several participants inquired about the differences between *ISPR* and the standard Qwerty layout after completing the experiment. The NASA-TLX scores for *ISPR* were also slightly lower than for *IDR* (see Figure 5e). This may be because the protective measures in *ISPR* are more inconspicuous.

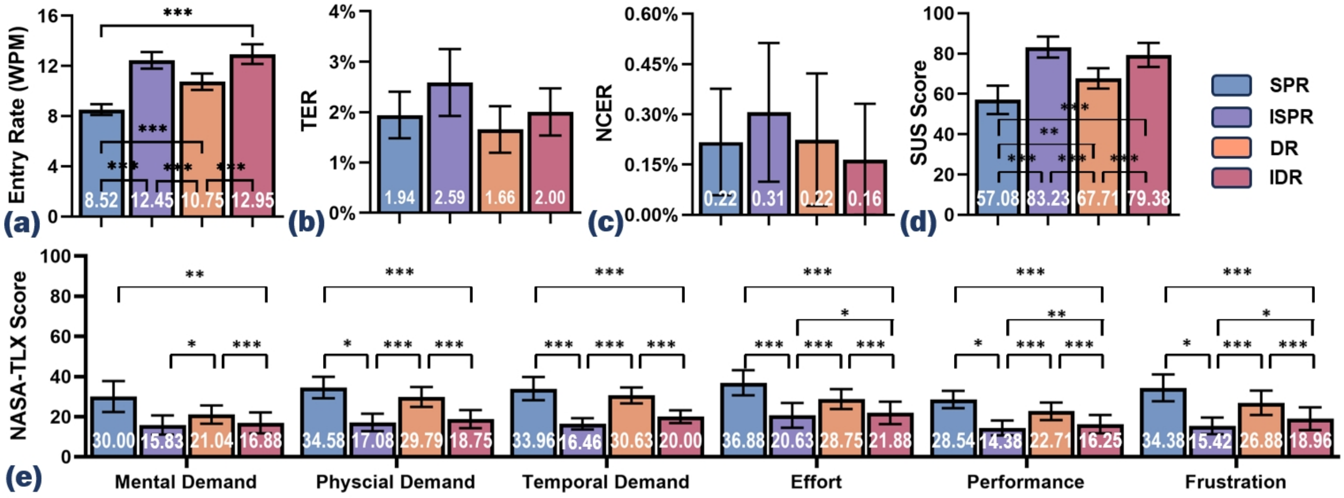

Figure 5: The means of (a) Entry Rate: (b) TER: (c) NCER: (d) SUS Scores (the higher, the better): and (e) NASA-TLX Scores (the lower, the better).of *SPR*, *ISPR*,*DR* and *IDR* in the user study.

Since our study primarily involved young adults, including both VR novice and experienced users, our findings can be generalized to this demographic. Our design relies on the standard Qwerty keyboard and straightforward password protection strategies, which are effective for both young and older adults [43]. Research by Wright et al. [55] indicates that keyboard preferences are not age-dependent. However, older adults often have decreased visual acuity [22], making it harder for them to track small cursor movements, especially with changes in ray direction.

### 6.5 Lessons

The following lessons (**L#**) were learned from this study.

**L3.** Randomly intermittently introducing variations in the virtual ray efficiently protects text content and achieves performance close to *STD* (**RQ.2**).

**L4.** When security mechanisms do not cause a significant decrease in typing performance, inconspicuous protective measures lead to a more positive user experience on text entry.

## 7 Design Recommendation for Protecting regular text Security (RQ.2)

- **Maintaining the Qwerty Configuration** When designing text security methods, it is advisable to maintain the Qwerty layout. Keeping the familiar Qwerty arrangement helps mitigate any potential declines in typing efficiency resulting from heightened security protocols.

- **Guaranteed Smooth Input Process** Smooth input is a key factor in enhancing text entry rates. It also could contribute to a positive subjective experience for users. When variations were introduced on the virtual ray at random intermittent frequency, both entry rates and subjective user experience were significantly improved in the user study (Section 6).

- **Invisible Encryption** Security measures should be designed to be imperceptible to users after ensuring that they do not affect the typing experience. Unobtrusive protection methods help enhance the subjective experience of users (**L5**).

## 8 Limitations and Future Work

This research has the following two limitations, which could serve as directions for future work. First, initially, our study employed leaked controller data to simulate potential text security breaches, leveraging its verified accuracy [56]. Nevertheless, unauthorized access to text content can also occur through covert surveillance methods such as clandestine photography. Subsequent research endeavours could broaden the scope of evaluation to encompass more diverse means of text acquisition and examine how introducing variations to virtual selection tools affects text security under these circumstances. Second, our study explored the implications of introducing variations on text security for text entry via controller-based raycasting, considering its wider applicability. Building upon the findings of this study, future research could extend this exploration to more alteration methods on virtual selection tools for different VR/AR typing techniques, such as drum-like typing [7] or finger tapping [17].

## 9 Conclusion

In this work, we studied the feasibility of introducing variations in virtual rays to enhance text security and the impact on text input performance with controller-based raycasting text entry technique in VR. By designing and evaluating approaches to introduce variability in the starting point and direction of virtual rays, with successive and random intermittent introduction forms, our goal is to enhance text security cost-effectively and ensure minimal disruption to usability. Through our evaluation, we found that introducing variations to virtual rays can enhance the security of regular text and passwords. The approaches of introducing variability with random intermittent frequency (*ISPR* and *IDR*) outperform successive introduction of variability (*SPR* and *DR*) in terms of typing performance of regular text. However, for passwords, which typically consist of a limited number of characters and are susceptible to forward inference, only a high-frequency introduction of variability to virtual rays is feasible.

As VR becomes more integrated into daily activities, securing text data in VR environments is crucial for protecting user privacy. Our method of introducing variability into virtual selection tools can be adapted in controller-based raycasting, offering a cost-effective solution for text security without extra effort. Ensuring text security is vital in virtual text input systems. Combining user authentication with text protection strengthens the defense against unauthorized access and interception of sensitive data.

## References

[1] Common password list ( rockyou.txt ). `https://www.kaggle.com/datasets/wjburns/common-password-list-rockyoutxt`.

[2] M. A. Bakar, Y.-T. Tsai, H.-H. Hsueh, and E. C. Li. Crowbarlimbs: A fatigue-reducing virtual reality text entry metaphor. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2806–2815, 2023. doi: 10.1109/TVCG.2023.3247060

[3] A. Bangor, P. T. Kortum, and J. T. Miller. An empirical evaluation of the system usability scale. *International Journal of Human–Computer Interaction*, 24(6):574–594, 2008. doi: 10.1080/10447310802205776

[4] X. Bi, B. A. Smith, and S. Zhai. Quasi-qwerty soft keyboard optimization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, p. 283–286. Association for Computing Machinery, New York, NY, USA, 2010. doi: 10.1145/1753326.1753367

[5] F. Binbeshr, M. Mat Kiah, L. Y. Por, and A. Zaidan. A systematic review of pin-entry methods resistant to shoulder-surfing attacks. *Computers & Security*, 101:102116, 2021. doi: 10.1016/j.cose.2020.102116

[6] C. Boletsis and S. Kongsvik. Controller-based text-input techniques for virtual reality: an empirical comparison. *International Journal of Virtual Reality (IJVR)*, 19(3), 2019. doi: 10.20870/IJVR.2019.19.3.2917

[7] C. Boletsis and S. Kongsvik. Text input in virtual reality: A preliminary evaluation of the drum-like vr keyboard. *Technologies*, 7(2), 2019. doi: 10.3390/technologies7020031

[8] L. Bošnjak and B. Brumen. Shoulder surfing experiments: A systematic literature review. *Computers & Security*, 99:102023, 2020. doi: 10.1016/j.cose.2020.102023

[9] J. Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[10] P. Casey, I. Baggili, and A. Yarramreddy. Immersive virtual reality attacks and the human joystick. *IEEE Trans. Dependable Secur. Comput.*, 18(2):550–562, mar 2021. doi: 10.1109/TDSC.2019.2907942

[11] S. Chen, J. Wang, S. Guerra, N. Mittal, and S. Prakkamakul. Exploring word-gesture text entry techniques in virtual reality. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, p. 1–6. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290607.3312762

[12] J. Dudley, H. Benko, D. Wigdor, and P. O. Kristensson. Performance envelopes of virtual keyboard text input strategies in virtual reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 289–300, 2019. doi: 10.1109/ISMAR.2019.00027

[13] J. J. Dudley, J. Zheng, A. Gupta, H. Benko, M. Longest, R. Wang, and P. O. Kristensson. Evaluating the performance of hand-based probabilistic text input methods on a mid-air virtual qwerty keyboard. *IEEE Transactions on Visualization and Computer Graphics*, 29(11):4567–4577, oct 2023. doi: 10.1109/TVCG.2023.3320238

[14] I. Flechais and M. A. Sasse. Stakeholder involvement, motivation, responsibility, communication: How to design usable security in e-science. *International Journal of Human-Computer Studies*, 67(4):281–296, 2009. doi: 10.1016/j.ijhcs.2007.10.002

[15] W. N. Francis and H. Kucera. Brown corpus manual. *Letters to the Editor*, 5(2):7, 1979.

[16] A. Giaretta. Security and privacy in virtual reality–a literature survey. *arXiv preprint arXiv:2205.00208*, 2022. doi: arXiv:2205.00208

[17] J. Grubert, L. Witzani, E. Ofek, M. Pahud, M. Kranz, and P. O. Kristensson. Text entry in immersive head-mounted display-based virtual reality using standard keyboards. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 159–166, 2018. doi: 10.1109/VR.2018.8446059

[18] 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

[19] A. Hinneburg. A density based algorithm for discovering clusters in large spatial databases with noise. *KDD Conference, 1996*, 1996.

[20] O. Hiroshi, H. T. Shinji Miyake, Masaharu Kumashiro, and K. Suzuki. Ranges of dynamic motion of the wrist in healthy young and middle-aged men. *Ergonomics*, 35(12):1467–1477, 1992. doi: 10.1080/00140139208967416

[21] R. S. Hirsch. Effects of standard versus alphabetical keyboard formats on typing performance. *Journal of Applied Psychology*, 54(6):484, 1970. doi: 10.1037/h0030143

[22] Z. Jia, P.-L. P. Rau, and G. Salvendy. Use and design of handheld computers for older adults: A review and appraisal. *International Journal of Human–Computer Interaction*, 28(12):799–826, 2012. doi: 10.1080/10447318.2012.668129

[23] H. Jiang and D. Weng. Hipad: Text entry for head-mounted displays using circular touchpad. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 692–703, 2020. doi: 10.1109/VR46266.2020.00092

[24] J. Kröger. Unexpected inferences from sensor data: A hidden privacy threat in the internet of things. In L. Strous and V. G. Cerf, eds., *Internet of Things. Information Processing in an Increasingly Connected World*, pp. 147–159. Springer International Publishing, Cham, 2019.

[25] J. Leng, L. Wang, X. Liu, X. Shi, and M. Wang. Efficient flower text entry in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 28(11):3662–3672, 2022. doi: 10.1109/TVCG.2022.3203101

[26] Y. Li, Y. Cheng, Y. Li, and R. H. Deng. What you see is not what you get: Leakage-resilient password entry schemes for smart glasses. In *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*, ASIA CCS '17, p. 327–333. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3052973.3053042

[27] Y. Li, Y. Cheng, W. Meng, Y. Li, and R. H. Deng. Designing leakage-resilient password entry on head-mounted smart wearable glass devices. *IEEE Transactions on Information Forensics and Security*, 16:307–321, 2021. doi: 10.1109/TIFS.2020.3013212

[28] Z. Ling, Z. Li, C. Chen, J. Luo, W. Yu, and X. Fu. I know what you enter on gear vr. In *2019 IEEE Conference on Communications and Network Security (CNS)*, pp. 241–249, 2019. doi: 10.1109/CNS.2019.8802674

[29] X. Liu, Z. Zhou, W. Diao, Z. Li, and K. Zhang. When good becomes evil: Keystroke inference with smartwatch. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, CCS '15, p. 1273–1285. Association for Computing Machinery, New York, NY, USA, 2015. doi: 10.1145/2810103.2813668

[30] X. Lu, D. Yu, H.-N. Liang, W. Xu, Y. Chen, X. Li, and K. Hasan. Exploration of hands-free text entry techniques for virtual reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 344–349, 2020. doi: 10.1109/ISMAR50242.2020.00061

[31] I. S. MacKenzie and S. X. Zhang. An empirical investigation of the novice experience with soft keyboards. *Behaviour & Information Technology*, 20(6):411–418, 2001. doi: 10.1080/01449290110089561

[32] A. Maiti, M. Jadliwala, and C. Weber. Preventing shoulder surfing using randomized augmented reality keyboards. In *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pp. 630–635, 2017. doi: 10.1109/PERCOMW.2017.7917636

[33] R. P. McMahan, C. Lai, and S. K. Pal. Interaction fidelity: The uncanny valley of virtual reality interactions. In S. Lackey and R. Shumaker, eds., *Virtual, Augmented and Mixed Reality*, pp. 59–70. Springer International Publishing, Cham, 2016.

[34] S. Millett. *Speed readings for ESL learners: 500 BNC*. School of Linguistics and Applied Language Studies, Victoria University of Wellington, 2017.

[35] E. Miluzzo, A. Varshavsky, S. Balakrishnan, and R. R. Choudhury. Tapprints: your finger taps have fingerprints. MobiSys '12, p. 323–336. Association for Computing Machinery, New York, NY, USA, 2012. doi: 10.1145/2307636.2307666

[36] M. Nabiyouni, A. Saktheeswaran, D. A. Bowman, and A. Karanth. Comparing the performance of natural, semi-natural, and non-natural locomotion techniques in virtual reality. In *2015 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 3–10, 2015. doi: 10.1109/3DUI.2015.7131717

[37] T. Ogitani, Y. Arahori, Y. Shinyama, and K. Gondow. Space saving text input method for head mounted display with virtual 12-key keyboard. In *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*, pp. 342–349, 2018. doi: 10.

1109/AINA.2018.00059

[38] E. Owusu, J. Han, S. Das, A. Perrig, and J. Zhang. Accessory: password inference using accelerometers on smartphones. In *Proceedings of the Twelfth Workshop on Mobile Computing Systems & Applications*, HotMobile '12. Association for Computing Machinery, New York, NY, USA, 2012. doi: 10.1145/2162081.2162095

[39] F. Rahutomo, T. Kitasuka, and M. Aritsugi. Semantic cosine similarity. 2012.

[40] V. Rajanna and J. P. Hansen. Gaze typing in virtual reality: Impact of keyboard design, selection method, and motion. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research &amp; Applications*, ETRA '18. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3204493.3204541

[41] D. Schneider, A. Otte, T. Gesslein, P. Gagel, B. Kuth, M. S. Damlakhi, O. Dietz, E. Ofek, M. Pahud, P. O. Kristensson, J. Müller, and J. Grubert. Reconviguration: Reconfiguring physical keyboards in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 25(11):3190–3201, 2019. doi: 10.1109/TVCG.2019.2932239

[42] D. Shukla and V. V. Phoha. Stealing passwords by observing hands movement. *IEEE Transactions on Information Forensics and Security*, 14(12):3086–3101, 2019. doi: 10.1109/TIFS.2019.2911171

[43] A. L. Smith and B. S. Chaparro. Smartphone text input method performance, usability, and preference with younger and older adults. *Human Factors*, 57(6):1015–1028, 2015. PMID: 25850116. doi: 10.1177/0018720815575644

[44] R. W. Soukoreff and I. S. MacKenzie. Metrics for text entry research: An evaluation of msd and kspc, and a new unified error metric. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, p. 113–120. Association for Computing Machinery, New York, NY, USA, 2003. doi: 10.1145/642611.642632

[45] M. Speicher, A. M. Feit, P. Ziegler, and A. Krüger. Selection-based text entry in virtual reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, p. 1–13. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3174221

[46] K. Vertanen and P. O. Kristensson. A versatile dataset for text entry evaluations based on genuine mobile emails. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, pp. 295–298, 2011.

[47] K. Viswanathan and A. Yazdinejad. Security considerations for virtual reality systems. *arXiv preprint arXiv:2201.02563*, 2022. doi: arXiv:2201.02563

[48] T. Wan, R. Shi, W. Xu, Y. Li, K. Atkinson, L. Yu, and H.-N. Liang. Hands-free multi-type character text entry in virtual reality. *Virtual Reality*, 28(1):8, 2024. doi: 10.1007/s10055-023-00902-z

[49] T. Wan, Y. Wei, R. Shi, J. Shen, P. O. Kristensson, K. Atkinson, and H.-N. Liang. Design and evaluation of controller-based raycasting methods for efficient alphanumeric and special character entry in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–11, 2024. doi: 10.1109/TVCG.2024.3349428

[50] T. Wan, L. Zhang, Y. Xu, Z. Guo, B. Gao, and H.-N. Liang. Analysis and design of efficient authentication techniques for password entry with the qwerty keyboard for vr environments. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–11, 2024.

[51] T. Wan, L. Zhang, H. Yang, P. Irani, L. Yu, and H.-N. Liang. Exploration of foot-based text entry techniques for virtual reality environments. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, CHI '24. Association for Computing Machinery, New York, NY, USA, 2024. doi: 10.1145/3613904.3642757

[52] C. Wang, X. Guo, Y. Chen, Y. Wang, and B. Liu. Personal pin leakage from wearable devices. *IEEE Transactions on Mobile Computing*, 17(3):646–660, 2018. doi: 10.1109/TMC.2017.2737533

[53] H. Wang, T. T.-T. Lai, and R. Roy Choudhury. Mole: Motion leaks through smartwatch sensors. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, MobiCom '15, p. 155–166. Association for Computing Machinery, New York, NY, USA, 2015. doi: 10.1145/2789168.2790121

[54] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. *The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures*, p. 143–146. Association for Computing Machinery, New York, NY, USA, 2011.

[55] P. Wright, C. Bartram, H. Emslie, J. Evans, B. Wilson, and S. Belt. Text entry on handheld computers by older users. *Ergonomics*, 43:702–16, 07 2000. doi: 10.1080/001401300404689

[56] Y. Wu, C. Shi, T. Zhang, P. Walker, J. Liu, N. Saxena, and Y. Chen. Privacy leakage via unrestricted motion-position sensors in the age of virtual reality: A study of snooping typed input on virtual keyboards. In *2023 IEEE Symposium on Security and Privacy (SP)*, pp. 3382–3398, 2023. doi: 10.1109/SP46215.2023.10179301

[57] W. Xu, H.-N. Liang, Y. Zhao, T. Zhang, D. Yu, and D. Monteiro. Ringtext: Dwell-free and hands-free text entry for mobile head-mounted displays using head motions. *IEEE Transactions on Visualization and Computer Graphics*, 25(5):1991–2001, 2019. doi: 10.1109/TVCG.2019.2898736

[58] Z. Xu, K. Bai, and S. Zhu. Taplogger: inferring user inputs on smartphone touchscreens using on-board motion sensors. In *Proceedings of the Fifth ACM Conference on Security and Privacy in Wireless and Mobile Networks*, WISEC '12, p. 113–124. Association for Computing Machinery, New York, NY, USA, 2012. doi: 10.1145/2185448.2185465

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

[60] H. Yamada. *A historical study of typewriters and typing methods, from the position of planning Japanese parallels*. Journal of Information Processing, 1980.

[61] C. Yildirim. Point and select: Effects of multimodal feedback on text entry performance in virtual reality. *International Journal of Human–Computer Interaction*, 39(19):3815–3829, 2023. doi: 10.1080/10447318.2022.2107330

[62] C. Yildirim. Point and select: Effects of multimodal feedback on text entry performance in virtual reality. *International Journal of Human–Computer Interaction*, 39(19):3815–3829, 2023. doi: 10.1080/10447318.2022.2107330

[63] C. Yildirim and E. Osborne. Text entry in virtual reality: A comparison of 2d and 3d keyboard layouts. In C. Stephanidis, J. Y. C. Chen, and G. Fragomeni, eds., *HCI International 2020 – Late Breaking Papers: Virtual and Augmented Reality*, pp. 450–460. Springer International Publishing, Cham, 2020. doi: 10.1007/978-3-030-59990-4_33

[64] M. R. Zhang, S. Zhai, and J. O. Wobbrock. Text entry throughput: Towards unifying speed and accuracy in a single performance metric. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, p. 1–13. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290605.3300866

[65] R. Zhang, N. Zhang, C. Du, W. Lou, Y. T. Hou, and Y. Kawamoto. Augauth: Shoulder-surfing resistant authentication for augmented reality. In *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6, 2017. doi: 10.1109/ICC.2017.7997251