Exploration of Foot-based Text Entry Techniques for Virtual Reality Environments

Tingjie Wan*
tingjie.wan20@student.xjtlu.edu.cn
Xi'an Jiaotong-Liverpool University
Suzhou, China

Liangyuting Zhang
liangyutingzhang@gmail.com
Xi'an Jiaotong-Liverpool University
Suzhou, China

Hongyu Yang
hongyu.yang22@student.xjtlu.edu.cn
Xi'an Jiaotong-Liverpool University
Suzhou, China

Pourang Irani
pourang.irani@ubc.ca
University of British Columbia
Kelowna, Canada

Lingyun yu
lingyun.yu@student.xjtlu.edu.cn
Xi'an Jiaotong-Liverpool University
Suzhou, China

Hai-Ning Liang†
haining.liang@xjtlu.edu.cn
Xi'an Jiaotong-Liverpool University
Suzhou, China

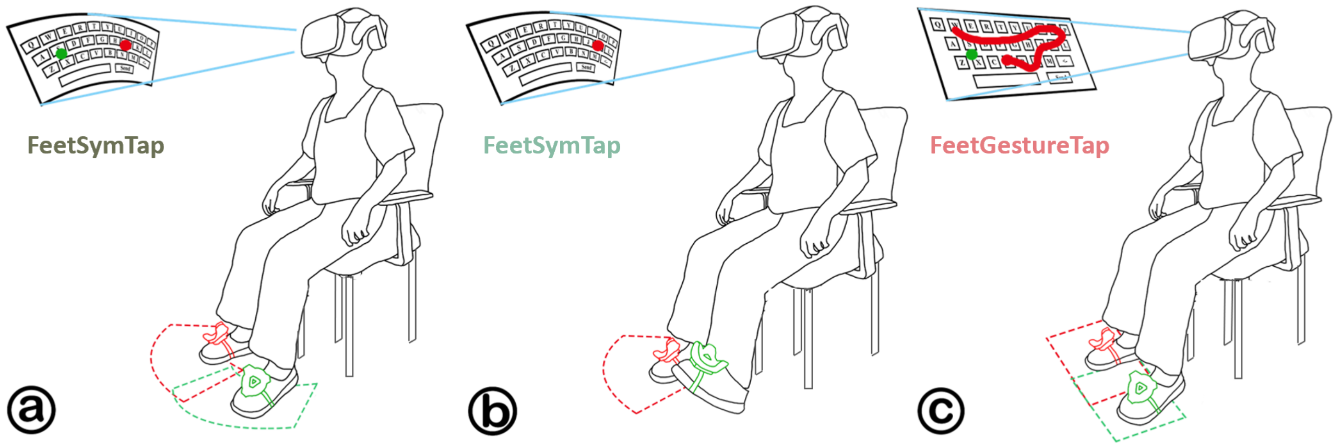

Figure 1: Text entry in a VR HMD using our three proposed foot-based techniques: (a) *FeetSymTap*, each foot controls one cursor to select target keys; (b) *FeetAsymTap*, right foot locates the position of the target and left foot makes the selection by toe tap; (c) *FeetGestureTap*, right foot inputs with the swipe-based method and left foot types with the tap-based method.

ABSTRACT

Foot-based input can serve as a supplementary or alternative approach to text entry in virtual reality (VR). This work explores the feasibility and design of foot-based techniques that are hands-free. We first conducted a preliminary study to assess foot-based text entry in standing and seated positions with tap and swipe input approaches. The findings showed that foot-based text input was feasible, with the possibility for performance and usability improvements. We then developed three foot-based techniques, including two tap-based techniques (FeetSymTap and FeetAsymTap) and one swipe-based technique (FeetGestureTap), and evaluated their performance via another user study. The results show that the two tap-based techniques supported entry rates of 11.12 WPM and 10.80 WPM, while the swipe-based technique led to 9.16 WPM. Our findings provide a solid foundation for the future design and implementation of foot-based text entry in VR and have the potential to be extended to MR and AR.

CCS CONCEPTS

- **Human-centered computing → Empirical studies in HCI**; **Virtual reality**; **Text input**.

KEYWORDS

virtual reality, text entry, foot-based interaction, hands-free interaction

ACM Reference Format:

Tingjie Wan, Liangyuting Zhang, Hongyu Yang, Pourang Irani, Lingyun yu, and Hai-Ning Liang. 2024. 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), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3613904.3642757

---

[*] Tingjie Wan is also with the Department of Computer Science, University of Liverpool, UK.

[†] Corresponding author.

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642757

1 INTRODUCTION

Virtual Reality (VR) has revolutionized our interaction with digital content, offering immersive experiences across various fields. With VR's evolution and wider adoption, the quest for efficient and natural text input methods has become increasingly important. While conventional handheld controllers and hand gestures have been the prevalent means for text input in VR, the limitations arise when the hands are preoccupied with other essential tasks or when using hands is impractical or socially challenging in certain scenarios, such as public venues.

Hands-free techniques have been investigated for text entry scenarios where users' hands are unavailable, including voice recognition, head-based input, and eye-gaze interaction. Voice recognition systems enable users to dictate text verbally; however, they may be affected by ambient noise and accuracy issues [16]. Head-based methods rely on users' head movements to select characters, which can increase neck fatigue and cause motion sickness [19, 32, 37, 38] or simulator sickness [54, 63, 67]. Eye-gaze interactions allow users to input text by fixating their gaze on the characters, but it may suffer from calibration issues, false-positive errors, and eye fatigue [44] and require a high level of focus and cognitive load [45].

Using the feet as an input source has been successfully explored in various use cases [61]. Foot-based interaction can be a viable and versatile approach for hands-free text entry. Foot typing, serving as an intuitive and hands-free method, allows users to interact with virtual keyboards using their feet. Foot-based typing offers a potential solution to issues linked with head-induced discomfort, like motion sickness or simulator sickness [19, 32, 37, 63, 67], often experienced in mobile settings such as traveling in a public bus or subway (see Figure 2a). Moreover, foot-based typing could free users' upper limbs. In seated activities like watching movies, it allows users to input text while keeping their upper limbs relaxed, mitigating the disruption in their comfort and immersion levels. It also can support interactions where the environment does not have a desk and where prolonged typing in via mid-air input can lead to hand and arm fatigue [22, 68] (see Figure 2b). The potential of foot-based techniques across various scenarios makes them an intriguing and increasingly relevant area of exploration.

To our knowledge, no prior research has explored hands-free text input in VR using foot-based methods. Therefore, the central contribution of our work lies in a comprehensive investigation of foot-based text entry in VR, focusing on hands-free techniques. In this work, we explored three relevant research questions (in Section 3.1) and then conducted a preliminary study (in Section 4) and one formal study (in Section 6) to explore and evaluate the design of the foot-based text entry techniques iteratively. Our exploration has led to seven lessons and five design considerations for foot-based text entry.

In short, our work presented in this paper makes the following contributions:

- We presented a first systematic exploration of the feasibility and applicability of foot-based text entry techniques for VR environments.
- We devised and evaluated three distinct foot-based techniques that demonstrated efficient performance with an acceptable workload.
- We introduced an arched Qwerty keyboard with an ergonomic layout designed to align with the natural movement trajectories of the feet and legs, enhancing user comfort and usability.

2 RELATED WORK

In this section, we first provide an overview of the landscape of foot-based interaction, concentrating particularly on its exploration in text entry. Then, we review the research about tap and swipe text entry methods, primarily within VR.

2.1 Foot-based Interaction

Foot-based interaction has become popular in HCI for two main reasons. First, the lower limbs offer a wide range of motion, enabling diverse and natural interactions. This range is influenced by joints such as the ankle, knee, and hip, as well as the user's posture [49]. For instance, in a sitting posture, the interaction range is confined to the area reachable by the feet, allowing for various foot and aerial (mid-air) gestures [52, 55, 60]. However, continuous lifting of both feet can lead to foot and leg fatigue. In a standing posture, only one foot can be used for interaction while the other stabilizes the body. During walking, the rhythmic, pattern-based movements of the feet restrict complex foot gestures for interaction, focusing primarily on balancing and covering the desired space [61]. Second, the availability of various devices capable of effectively capturing foot-related information has made it feasible to implement and explore foot-based interaction methods. These devices include foot-worn sensors, such as pressure sensors (e.g., used in [52, 59]) or inertial measurement units (IMUs) (e.g., used in [47]), which can precisely track the movement and orientation of the feet in real-time [61, 77]. Moreover, VR HMDs often come equipped with motion-tracking systems that can detect foot movements when combined with additional foot-tracking devices like the Vive Trackers, Kinect sensors, Vicon motion tracking system [51, 52, 77]. This integration enables users to interact with virtual environments using their feet, enhancing the sense of immersion and naturalness during VR interaction.

Some prior research has demonstrated that foot interaction can be used for precise selection [24, 46, 47]. Saunders and Vogel [51] investigated indirect foot pointing through discrete taps and kicks while standing. They found a preference for toe taps during interaction, followed by whole foot taps and then heel taps. Additionally, users in a standing position showed the ability to alternate between both feet. Felberbaum and Lanir [14] explored foot interactions across different stances and identified symmetric foot postures and a preference for the dominant right foot among users. These insights inspired our exploration of foot-based text input in VR, as text entry involves a sequence of pointing and confirmation selection actions. Saunders and Vogel [52] further investigated indirect foot interaction techniques for standing desks, emphasizing the role of foot postures not only in confirmation selections but also in pointing toward targets. However, text entry in VR presents a more intricate challenge, demanding high-frequency and precise pointing and selection actions due to densely arranged targets like

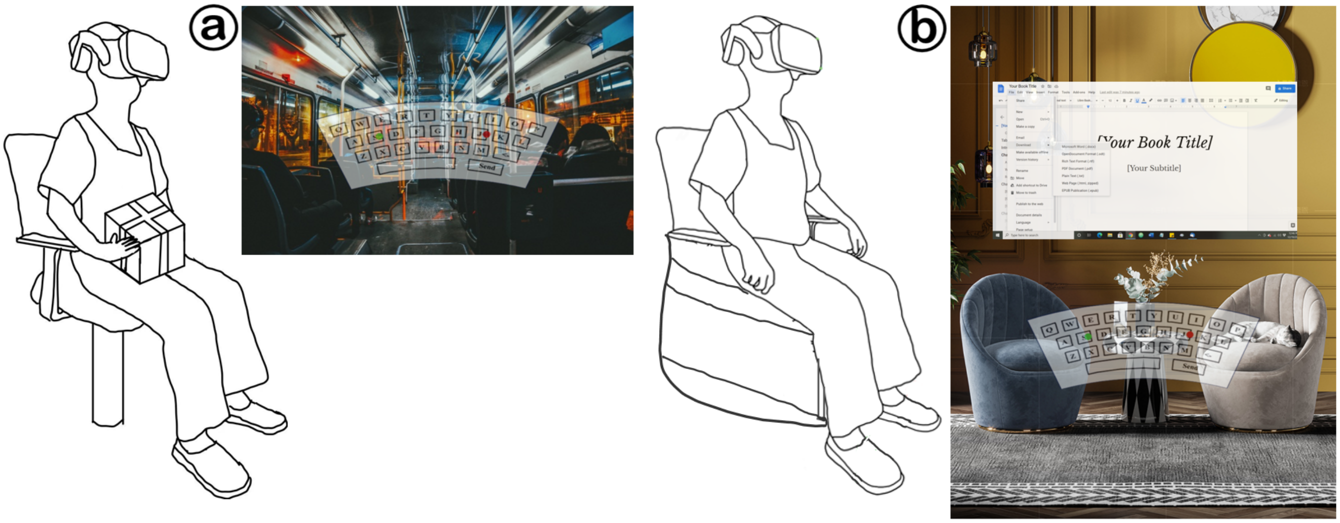

Figure 2: Two examples of potential scenarios for foot-based text entry. (a) A user is sitting on a bus or subway and holding something with both hands but wants to search for information on a website. Head typing could result in motion sickness, so he could type with one foot or both feet. (b) A user is writing a document in a hotel room, but there is no desk around, and the long period of mid-air typing makes is tiring, so the user alternates hands and feet to edit the text.

characters on a keyboard. Users continuously locate and select different characters, requiring greater precision and subtle movements for accurate and error-free input.

Exploring foot-based interaction for text entry addresses two primary motivations. The first is its potential to serve as an additional input method to enhance efficiency and reduce workload [47]. For instance, Rajanna et al. [47] incorporated foot gestures as a confirmation method for gaze-based text entry, resulting in improved text entry rates (foot gesture: 14.98 WPM, and foot press: 13.82 WPM) and reduced visual fatigue. The second motivation involves providing an alternative input modality, particularly when users' hands are occupied with other tasks or unsuitable for use. Some attempts have been made to enable text entry using foot gestures, but they only achieved relatively low text entry rates. Dobosz and Trzcionkowski [10] explored typing with four-foot gestures (toe tap, heel tap, toe rotation, and heel rotation) detected by a Myo armband but achieved only 1.46 WPM with IMU and 1.23 WPM with EMG. Pedrosa et al. [42] developed the DuoGrapher and SwingingFoot typing methods based on heel rotation for individuals with motor impairments, but both methods yielded text entry rates below 5 WPM. Tao et al. [59] designed a shoe keyboard for typing through insoles that detect pressure and acceleration. While they analyzed the time cost of typing a character to be approximately 2.87 seconds, they did not run user studies to determine the actual text entry rates that users could achieve.

These efforts highlight the challenges and potential of foot-based text entry in various contexts. As stated earlier, to our knowledge, no prior work has systematically examined the feasibility and applicability of text entry via foot-based techniques for VR scenarios. Our work thus aims to fill this gap.

2.2 Text Entry with Tap

The most well-known virtual keyboard relies on tapping for key selection. It became ubiquitous in modern touchscreen devices (e.g., smartphones), and its use has extended into VR. Tap metaphors have been extensively explored and compared in various studies, including pointing to keys using ray casting with handheld controllers or head motions captured from the HMD (e.g., [5, 58, 64, 65, 75]) and touching virtual keys with hand/hand gestures (e.g., [2, 5, 12, 40, 58]).

In VR, the effectiveness of tap-typing techniques has been evaluated using different input methods, including controller-based touching and pointing, head-pointing, and even hand gestures. Studies have consistently shown that controller-based methods outperform other approaches regarding text entry rate. For instance, Boletisis et al. [5] found that controller-based touching (21.01 WPM) achieved the highest text entry rate in VR, followed by controller pointing (16.65 WPM), head-pointing (10.83 WPM), and trackpad-based selection on a split keyboard (10.17 WPM). Similarly, Spericher et al. [58] conducted an evaluation of VR text entry techniques and identified controller pointing as the most effective method (15.44 WPM), followed by controller contact tapping (12.69 WPM). Head pointing (10.20 WPM) also outperformed two-handed touch tapping (9.77 WPM) in some cases [5]. Additionally, Xu et al. [70] compared pointing with controllers (14.6 WPM) and pointing with the head (5.62 WPM) in AR and presented findings consistent with those in VR.

Researchers have explored various tap-typing configurations, including one-handed and two-handed typing, split keyboard layouts, and other approaches like using the palm of the non-dominant hand as an interactive keyboard and the dominant hand to tap on the palm of the non-dominant hand [66]. Adhikary and Vertanen [2] investigated contact tap-typing VR text entry with one or two hands, considering both traditional Qwerty layout and split keyboard layout. They observed no significant difference in objective performance and workload between one-handed typing (16.1 WPM) and two-handed typing (16.4 WPM). However, they found that the text entry performance of the split layout with two hands was lower compared to the traditional Qwerty layout. On the other hand, the study conducted by Rickel et al. [48] involving one-handed and two-handed mid-air tap input methods with the Microsoft HoloLens 2 revealed that typing with two hands was faster and more preferred than typing with one hand. Nevertheless, the two-handed tap input method was also found to be less accurate and associated with higher fatigue in certain body parts.

In previous work, researchers have evaluated VR tap-typing techniques. The superior performance of controller pointing compared to bimanual touch tapping and the potentiality of some innovative approaches both would suggest that direct adaptation from traditional input methods and how text entry is done in the physical environment into VR scenarios do not always bring reciprocal performance. As such, it is also important to determine how applicable tap-typing via users' feet is in VR, as this is unexplored.

2.3 Swipe-based Keyboard

Word-gesture keyboards are commonly implemented using swipe motions and were initially designed for handwriting and stylus input [27, 78]. These keyboards allow users to input text by tracing continuous gestures on virtual keyboards, predicting words based on the gestures created. They have found widespread use in various contexts, and have been adopted in mobile touchscreen devices (e.g., [8, 9, 28, 53]), real-world mid-air interaction [3, 36], VR [6, 11, 21, 23, 68, 70, 72, 74, 75], and AR [33]. While most word-gesture keyboards are designed for one-handed use, bimanual gesture keyboards [4], involving two-handed input on a split Qwerty keyboard, are also available, but they did not exhibit superior performance.

Regarding VR text entry, the choice between swipe and traditional tap-typing keyboards depends on several contextual factors. For instance, when typing with a smartphone for a VR environment, studies have shown no significant performance differences between tap-typing and word-gesture keyboards [21]. However, controller-based word-gesture keyboards for VR are less user-friendly and efficient than tap-typing keyboards with controllers [23]. Word-gesture head-based typing is often seen in AR HMDs and, in some instances, can lead to a faster text entry rate compared to tapping with a physical button or using dwell-based techniques [33, 75].

Different word-gesture input methods have been explored, including using controllers [6], mid-air hands or finger gestures [11, 18, 68], head pointing [33, 70, 72]. These approaches achieved varying text entry rates, error rates, and user experiences, suggesting that the design and implementation of word-gesture keyboards can play a crucial role in their success.

In summary, swipe keyboards have demonstrated their potential to enhance text entry performance and experiences across various platforms, including VR.

3 RESEARCH QUESTIONS, APPROACH, AND IMPLEMENTATION

In this section, we define research questions and our approach to answering them and then introduce the implementation of the foot-based text entry techniques.

3.1 Research Questions and Approach

We formulated the following three research questions (**RQ#**) to systematically explore the feasibility of using users' feet to achieve an acceptable text input performance and user experience.

- **RQ1**. Is foot-based text entry feasible when sitting and standing in VR scenarios? Does it require long-term learning?
- **RQ2**. How do we design foot-based text entry techniques for VR to achieve an acceptable hands-free text entry performance with a light workload?
- **RQ3**. Is foot gesture-based typing better than foot tap-based typing, and vice versa? What levels of text entry and error rates can these two typing methods achieve?

To answer these RQs, we first described our implementations that cover four design considerations: applicability, learnability, efficiency and accuracy, and fatigue minimization. Our exploration started with a preliminary study to assess the feasibility and learnability in both sitting and standing postures (**RQ1**). The gained insights were instrumental in refining the design of foot-tap and swipe-based text entry techniques. In the next user study, we compared the performance and user experience of three techniques (*FeetSymTap FeetAsymTap*, and *FeetGestureTap*) designed based on the results of the preliminary study. The results from the preliminary study and the user study would help answer **RQ2** and **RQ3**.

3.2 Implementation

3.2.1 *Apparatus.*

We used an HTC Vive Pro 2 for this experiment. It had a dual RGB low persistence LCD screen, a 2448 × 2448 pixels per eye resolution, and a 120Hz refresh rate. It was connected to a Windows 10 Pro PC with an Intel i9-11900 CPU and an Nvidia GeForce GTX 3090 GPU. The techniques and virtual environment were implemented using Unity3D (v2021.3.1f1) with SteamVR Unity plugin (version 2.7.3) and an HTC Vive Tracker 2.0. We chose a comparatively low-cost setup to ensure the wider **applicability** of our findings to devices available in the current market.

3.2.2 *Typing Interface.*

Users' reluctance to learn new layouts, as adapting to different layouts can be challenging for them [2], have contributed to the frequent adoption of Qwerty keyboard layout in VR virtual keyboards [7]. Thus, we used the Qwerty keyboard layout to ensure low **learnability** costs.

In the VR environment, we moved the keyboard interface from under the user's feet to their direct line of sight. This change aims to improve user comfort and efficiency by eliminating the need for frequent view rotations to see the keyboard under users' feet [16, 39, 51]. The interface in the VR HMD consists of a text display area and a virtual keyboard. The text display area shows both the transcribed sentences and the input entered by users. The virtual keyboard is positioned 10m in front of users, and its size is set to 3.6m × 1.4m, as Figure 3 shows. All character keys had the same size (0.3m × 0.3m). These features are similar to those in prior work dealing with text entry techniques.

3.2.3 Foot-based Interaction. The range of motion of the human knee and hip joints is generally similar, although individual reach ranges may vary due to differences in leg lengths. To minimize potential uncertainties, we selected participants with specific heights, including two females with heights of 171cm and 162cm and two males with heights of 186cm and 175cm. This allowed us to determine the size of the foot-operated keyboard as 80cm × 35cm, ensuring that participants could comfortably move their feet within this range.

As discussed in Sections 2.2 and 2.3, tap-based and swipe-based approaches have been shown to enable highly efficient and accurate text entry in VR. We mainly focused on foot-based text entry methods using these two interaction approaches given their superior **efficiency and accuracy**. There are several foot gestures that could possibly support these two interaction approaches. Toe tap and heel tap gestures are commonly used foot gestures for selecting options or activating specific functions. The toe tap is particularly popular among these gestures due to its ergonomic advantages because it involves minimal effort and primarily relies on the movement of the ankle joint, making it a **fatigue-minimizing** choice for foot-based interactions [52, 61]. Based on several pre-pilot trials, we confirmed that HTC Vive Trackers could recognize foot-tap actions accurately.

Pointing with the feet is primarily performed with the feet resting on the floor, as continuously floating the feet can quickly lead to leg fatigue, especially when standing because standing on one foot is challenging [61]. The feet rest naturally on the floor most of the time, while the feet may lift unintentionally during movement. To differentiate intentional toe tap for selection purposes from natural foot-tip elevations, we established a criterion: selection activation occurs when the HTC VIVE Trackers strapped to the toes have an upward lift of more than 10 degrees. This value was chosen based on the average dorsiflexion range [61] and our pre-testing. We identified user toe tap by calculating the distance and direction of movement from the HTC VIVE Trackers. In other words, during typing, users can lift their feet moderately, as long as the elevation does not surpass the threshold used to recognize toe tap.

3.2.4 Statistical Decoding Algorithm. The tapping-based text entry method handles the noise of input and predicts the input words with a statistical decoding algorithm [15]. The basic principle of the statistical decoding algorithm is to use maximum likelihood estimation and Bayesian inference to select the most likely words based on the entered characters and statistical information on word frequency and probability in the lexicon. We employed a lexicon of 10K words, which consisted of the most probable words extracted from the American National Corpus [1].

3.2.5 Word-Gesture Recognition. We implemented the gesture-word recognition algorithm by referencing existing works like SHARK2 [27]. The word gesture system decodes gestures into predefined words from the lexicon mentioned in Section 3.2.4. Initially, each word in the lexicon is converted into a line connecting the key center points of sequential letters in the word. When processing a gesture, the algorithm starts by filtering out words in the lexicon whose start/end locations are more than one key-width away from the start/end of the gesture. Subsequently, the algorithm calculates the distance between a candidate word and the input gesture by the sum of the Euler distances.

4 PRELIMINARY STUDY: EXPLORING FOOT-BASED TEXT ENTRY WHEN STANDING AND SITTING

This study evaluated the feasibility of foot-based text entry in VR in standing and sitting postures. We devised two distinct one-foot text entry techniques, *FootTap* and *FootGesture*, as shown in Figure 3.

4.1 Single-foot Text Entry Techniques

4.1.1 FootTap. FootTap is a discrete text entry method (the left keyboard in Figure 3). The user can use their non-support foot, which does not provide the primary support for the body, to interact with a virtual keyboard displayed in the VR environment. The user can move their non-support foot on the floor to select the desired keys on the virtual keyboard. Once the desired key is highlighted or targeted, the user can then confirm the selection by performing a toe tap with the same non-support foot.

4.1.2 FootGesture. FootGesture is a word-gesture text entry method adapted from touch-based word-gesture keyboards (the right keyboard in Figure 3). In this method, the user initiates the gesture by a toe tap of the non-support foot, indicating the start of the gesture. The gesture shape of the word is then traced when the non-support foot slides on the floor, and the gesture is concluded by another toe tap, indicating the end of the gesture. After, the user selects the target word from the candidate words by moving their non-support foot and toe-tapping it. This method leverages the foot's natural dexterity and movement capabilities to enable seamless text entry in VR. The foot's motion allows for a smooth and precise drawing of the word's gesture, which is then interpreted by the system.

4.2 Participants

We recruited 8 participants (4 males and 4 females; aged between 19 to 25, $M = 22.63, SD = 1.93$; heights from 160cm to 188cm, $M = 170.11, SD = 10.33$) from a local university. All participants were familiar with the Qwerty layout. participants had experience with VR before.

4.3 Experiment Design and Procedure

This study employed a within-subjects design with Technique and Posture as the two independent variables, resulting in four conditions. The order of the four conditions was counterbalanced using a Latin-Square approach. In each condition, participants transcribed 12 sentences sourced from MacKenzie and Soukoreff's phrase set [35]. The sentences were randomly selected without duplicates. The first two sentences were designated for training purposes, and participants' performance on these sentences was not recorded. The subsequent ten sentences were considered formal trials and were recorded for analysis. A total of 320 trials were used for analysis (= 8 participants × 2 text entry techniques × 2 postures × 10 sentences).

At the start of the experiment, participants completed a consent form and a demographics questionnaire. They were then introduced to the VR device, tasks, and techniques. After, participants were

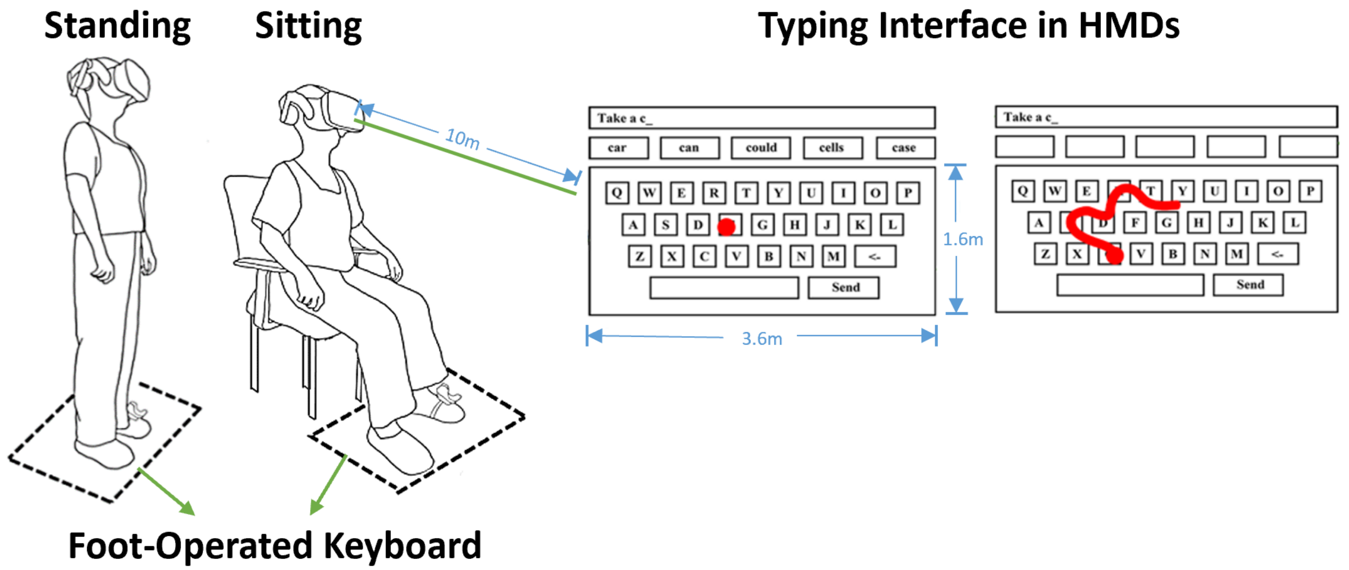

Figure 3: The user typing with *FootTap* (the left keyboard) and *FootGesture* (the right keyboard) in seated and standing positions. The two keyboards are the typing interfaces in the VR HMDs, and the foot-operated keyboard on the floor is a rectangle Qwerty layout but is invisible in the real world.

asked to wear the headset and begin the experiment. The experiment comprised four sessions, each involving one of the techniques and one of the postures. Participants were asked to complete the task primarily focusing on speed and accuracy. To prevent the influence of the chair on the range of motion of the feet, participants were asked to sit in a stationary chair in each session. Participants were asked to complete post-task questionnaires after each session and participate in semi-structured interviews to gather their feedback and suggestions after finishing four sessions. A five-minute break was provided between two sessions, with additional time given if requested by a participant. The entire experiment lasted approximately 60 minutes.

4.4 Evaluation Metrics

We measured the performance of the two text entry techniques for each posture using the objective data recorded during the experiments. Additionally, we gathered subjective feedback through three questionnaires.

- **Entry Rate** was measured in words per minute (WPM) [73]. This metric was computed by taking the number of transcribed words and dividing it by the time it took to complete the text transcription, which was measured in minutes. A word was defined as a continuous sequence of five characters, including spaces.
- **Error Rate** [57] was determined using standard word-level typing metrics, with the total error rate (TER) being the sum of the not corrected error rate (NCER) and the corrected error rate (CER).
- **Workload** associated with text entry methods and postures was evaluated using the NASA-TLX workload questionnaire [20]. This questionnaire consists of six subscales representing different aspects of workload, including mental demand, physical demand, temporal demand, frustration, effort, and performance. Participants rated each subscale on a scale ranging from 0 to 100, with intervals of 5. Lower scores indicate a lower workload and better overall performance.
- **Usability** was assessed using two different questionnaires: the After Scenario Questionnaire (ASQ) [30] and the Post-Study System Usability Questionnaire (PSSUQ) [31]. ASQ is a scenario-based usability questionnaire comprising three statements designed to assess the usability of text entry methods under each posture. PSSUQ is used to evaluate the overall system usability of each text entry method. PSSUQ includes three sub-scales: system usefulness, information quality, and interface quality. For ASQ and PSSUQ, participants provided ratings for each statement on a 7-point Likert scale, with a rating of 7 indicating "Strongly disagree" and a rating of 1 indicating "Strongly agree." Participants also had the option to mark the prompts as "N/A" (not applicable). The overall result for both ASQ and PSSUQ was computed by averaging the scores across the seven points of the scale.
- **Interview** was a semi-structured regarding (1) participants' willingness to type with their foot when hands are unavailable; (2) the experience of *FootTap* and *FootGesture* when used in each posture; and (3) any possible improvements for foot-based text entry.

4.5 Results and Discussion

We used SPSS 26 for data analysis. Shapiro-Wilk tests and Q-Q plots indicated that entry rate, ASQ, PSSUQ, and NASA-TLX data were normally distributed ($p > .05$), while TER and NCER were not normally distributed ($p < .05$). Thus, we applied Aligned Rank Transform [69] to TER and NCER 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.

4.5.1 Entry Rate and Error Rate. An RM-ANOVA analysis indicated that Posture had a significant main effect on entry rate ($F_{1,7} = 9.292, p = .019, \eta_p^2 = .670$), while Technique did not show a significant difference ($p > .05$), as illustrated in Figure 4a.

The RM-ANOVAs demonstrated a statistically significant effect of Technique on TER ($F_{1,7} = 16.599, p = .005, \eta_p^2 = .703$), but no significant differences in NCER ($p > .05$). Further, they did not reveal any significant differences in Posture for TER and NCER ($p > .05$) (see Figure 4b and c).

In general, *FootTap* and *FootGesture* both achieved acceptable entry rates (both over 7 WPM when sitting and 6 WPM when standing) and error rates (below 0.1%) in this short-term study. In the standing posture, *FootTap* and *FootGesture* led to slower performance when sitting (6.73 WPM with *FootTap*; 6.67 WPM with *FootGesture*).

4.5.2 Usability and Perceived Workload. Figures 5a and b summarize the ASQ and PSSUQ results, respectively. An RM-ANOVA found a significant main effect of Posture in ASQ scores ($F_{1,7} = 5.060, p = .050, \eta_p^2 = .445$). We did not find any significant differences in PSSUQ scores between the two techniques ($p > .05$).

Figure 5c summarizes the mean NASA-TLX scores of each technique under each posture. Test results show that Posture had a significant effect on workload across all six dimensions ($F = 246.524, p = .004, Wilks'\Lambda = .001, \eta_p^2 = .999$). For each dimension of NASA-TLX, RM-ANOVAs showed significant effects of Posture in mental demand ($F_{1,7} = 5.948, p = .045, \eta_p^2 = .459$), physical demand ($F_{1,7} = 73.085, p < .001, \eta_p^2 = .913$), temporal demand ($F_{1,7} = 21.415, p = .002, \eta_p^2 = .754$), effort ($F_{1,7} = 7.498, p = .029, \eta_p^2 = .517$), and frustration ($F_{1,7} = 11.641, p = .011, \eta_p^2 = .624$). Technique did not lead to any significant differences in workload.

Similar to the results in entry and error rates, ASQ and NASA-TLX scores for the two techniques were significantly higher in the standing condition than in the seated position, which means users prefer to use their feet to type while seated because their workload (mainly physical demands) is lower (see Figures 5a and c).

4.5.3 Interview. All participants (N=8) unanimously agreed that foot-based text input is feasible when hands are unavailable. However, they demonstrated a preference for using foot typing for lightweight tasks, such as sending instant messages. Particularly in a seated posture, with feet readily available on the ground, no adjustment in position is required. For more extensive typing tasks, while foot typing may not sustain prolonged use, it can serve as an alternating text input technique alongside hand typing, especially when physical keyboards are not utilized. Additionally, when the head movement is limited, or the hands are preoccupied with other essential tasks, foot-based typing emerges as a promising alternative, potentially enabling effective communication.

Users exhibited a distinct preference for the right foot (dominant foot) when opting for foot input, regardless of whether in a standing or seated position. Without exception, all users selected the right foot for input, with the left foot primarily supporting in the standing position and assisting in posture adjustments while seated, analogous to the stabilization role of the left hand [17]. This is consistent with the principle of left-hand priority in Guiard's kinematic chain model of asymmetric bimanual tasks [17], where both the left hand and left foot handle tasks contribute less precision or positioning adjustments, while the right hand and right foot undertake actions requiring higher precision and finer motor control.

Participants expressed that *FootTap* was easier to grasp, and although they learned to use *FootGesture* quickly, they still needed some training to become proficient since they did not often work with word-gesture-based typing techniques. They acknowledged that typing while standing was more physically demanding than seated, with seven participants mentioning that *FootTap* was the most tiring and only one participant (P8) reporting *FootGesture* as more exhausting. Typing with one foot while standing requires the other foot to bear the balance of the entire body, potentially causing leg fatigue. Even when using both techniques in a seated posture, six participants still experienced some leg fatigue because the knee had to rotate when positioning characters, and some distant characters required large-amplitude leg movements.

Regarding *FootTap*, all participants expected to be able to utilize both feet similarly to how they would use their hands for typing with a physical keyword, which could reduce leg fatigue. However, using both feet simultaneously while standing was not feasible due to the need for one leg to help maintain balance. Furthermore, four participants also had an expectation that the confirmation of target selection could be more precise, as one foot was tasked with both locating target keys and confirming the selection of the location, which increased motor coordination requirements. Although doing both tasks with one foot gave them some challenges, they still thought that the confirmation of the toe tap was better compared to other foot confirmation methods, such as the heel lift. This was because the toe lift would be a labor-saving action that people could easily do.

On the other hand, in the case of *FootGesture*, five participants felt limited by the requirement to make word-level entries, hindering their flexibility in typing. For instance, when typing the common word 'the', a common short word, using *FootTap*, they only needed to tap the first letter 't' to access the word from the predictions. In contrast, *FootGesture* required the complete typing of the word, which posed a potential drawback.

4.6 Lessons Learned from the Preliminary Study

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

- **L1.** Users could enter text with tap- and swipe-based typing methods in seated and standing positions without requiring extensive learning. However, it is less feasible or practical to perform foot typing in a standing position for a long time due to physical fatigue (**RQ1**).

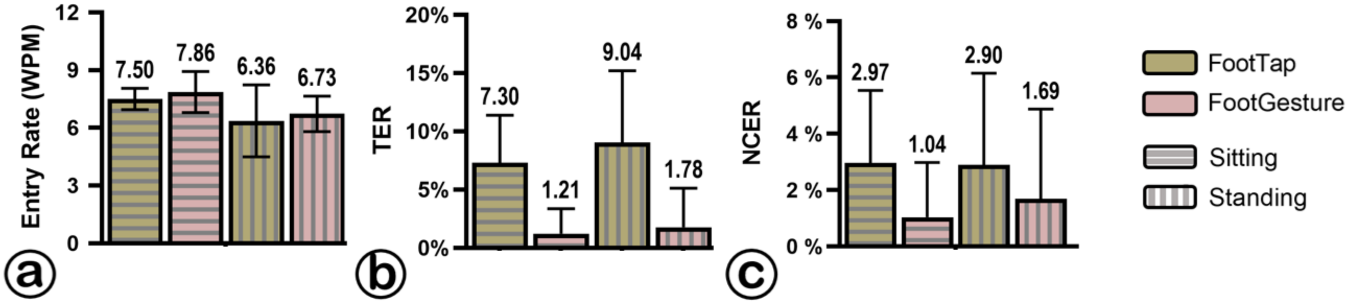

Figure 4: The means of (a) Entry rate, (b) TER, and (c) NCER.

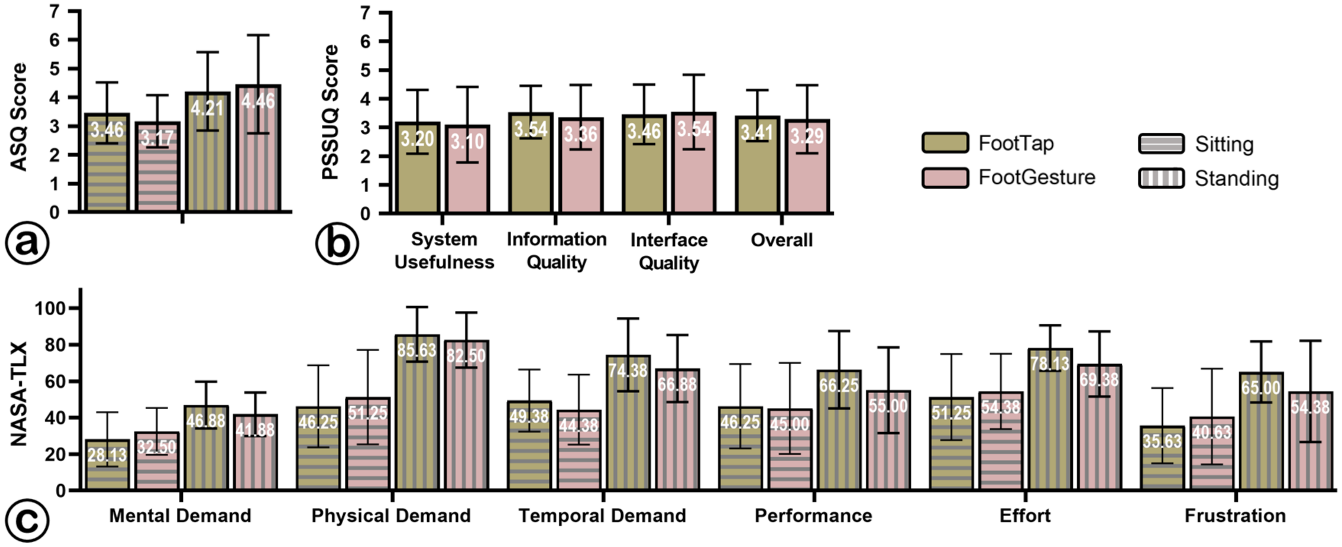

Figure 5: The means of (a) ASQ, (b) PSSUQ, and (c) NASA-TLX scores.

L2. Short, simple text entry tasks are more appropriate when standing, such as entering passwords and sending short messages (e.g., using an instant messaging app), especially using *FootGesture*. In the seated position, lightweight text entry tasks were more feasible, as there was no need for the legs and feet to help maintain body balance.

L3. Compared to typing via only one foot, participants were more expecting to coordinate and utilize both feet for text entry, which may help to reduce leg fatigue and improve the precision of selection confirmation.

L4. The swipe-based approach (*FootGesture*) tends to involve less fatigue compared to the tap-based approach (*FootTap*) in standing posture. However, the swipe-based approach's inflexibility may hinder its use in certain cases, such as typing some short words.

5 USER-INSPIRED BIPED-BASED TEXT ENTRY TECHNIQUES

Based on the findings from the preliminary study, we improved our foot-based text entry techniques from the following three aspects: (1) shifting from uniped-based to biped-based text entry, (2) modifying the rectangular layout to arch-shaped for tap-based text entry techniques to reduce the leg fatigue, and (3) integrating a tap mechanism into the swipe-based text entry approach to solve its inflexibility. Before introducing the proposed techniques, we first describe the design rationales for the arch-shaped keyboard layout.

Changing the Qwerty keyboard to an arch-shaped one is primarily motivated by the physiological structure and movement patterns of users' feet, which is supported by Müller et al.'s design [39] that implemented a semi-circular interface when exploring foot-operated user interfaces and Wenge et al.'s design [71] which used a distorted circular interface to support directional foot movements for selecting targets from a fixed standing position. The human foot is composed of the heel and the toes. When the heel remains stationary, the motion of the toes naturally follows an arched trajectory and primarily involves the ankle joint. In contrast, if the entire foot or the heel is in motion, both the ankle and knee joints are engaged. This design considers that users tend to rotate their feet around their heels rather than dragging them across the floor, as also observed in Velloso et al.'s studies [60]. This rotation minimizes the involvement of knee joints and horizontal foot movements. Designing the keyboard in an arch form minimizes the rotation of the knee joints and horizontal foot movement, thus reducing leg fatigue. This keyboard design capitalizes on the anatomical compatibility of the arch shape with the foot's inherent movement tendencies, promoting more ergonomic and comfortable interactions, as Figure 6a shows. Using an arched keyboard, users can simply pivot the ankle joint to reach keys close to the current cursor position (see Figure 6b). For keys situated at a greater distance, a combination of ankle joint rotation and slight knee rotation can be employed to reach the desired keys' location (see Figure 6c). In the foot-operated keyboard, the arc length of the upper arc is 65cm, the arc length of the lower arc is 60cm, and the distance from the upper arc to the lower arc is 27cm (see Figure 7a). The movement range of the HTC VIVE Tracker is the same as the size of the foot-operated keyboard.

The decision to use the arched keyboard exclusively for tap-based techniques and not for swipe-based techniques stems from the nature of the two methods and ergonomic considerations. Tap-based techniques involve single-character input, where users select one character at a time. The arched keyboard is well-suited for this method because it allows efficient tapping with minimal leg displacement. Users can pivot their ankles to reach keys, reducing the need for extensive leg movement. Conversely, swipe-based techniques require the sequential input of multiple characters through swiping gestures. A rectangle keyboard layout is more suitable for swiping as it encourages relatively straight trajectories for swiping gestures. This layout enhances the fluidity and coherence of swiping gestures by minimizing unnecessary changes in direction between successive swipes.

We next introduce three biped-based typing techniques, including two tap-based techniques (*FeetSymTap* and *FeetAsymTap*) and one word-gesture-based technique (*FeetGestureTap*).

5.1 FeetSymTap

FeetSymTap is a bipedal discrete symmetric text entry technique that provides users with two cursors, each representing one foot. These cursors select characters from a virtual Qwerty keyboard through foot movements. The character selection is confirmed by performing a toe tap on the foot corresponding to the active cursor. The toe tap action requires users to raise and lower their toes. This design is adapted from typing with both hands, allowing the user to type more fluidly and coordinately and helping prevent overuse or strain on a single foot. Symmetry in foot gestures is a pivotal aspect of *FeetSymTap*, considering the inherent coordination and simultaneous use of both feet during interactions [14]. Saunders et al.'s discovery [51] that the ability to perform discrete clicks with both feet while standing shows minimal dominance between feet further supports the importance of symmetric foot gestures.

When typing with both hands, each hand typically assumes a distinct role—typically, the left hand controls the left side of the keyboard while the right hand manages the right side. However, people's feet lack the same level of dexterity as their hands. In daily activities, humans often alternate the use of both feet. Additionally, Felberbaum et al. [14] identified a user preference for the right foot. This suggests that during usage, users may naturally alternate between both feet, with a tendency to favor the right foot more. Typing on a keyboard does not allocate the left foot solely for inputting characters on the left side and the right foot for the right side. Consequently, situations may arise where one foot attempts to access keys closer to the other foot. Due to the limited flexibility of the feet, one foot might not promptly yield space to accommodate the movement of the other foot. Consequently, this obstruction could impede the movement or access of the other foot during typing. To avoid one foot obstructing another during typing, the positions of the foot-operated keyboards under two feet are intentionally misaligned, as shown in Figure 7a.

5.2 FeetAsymTap

FeetAsymTap is an asymmetric bipedal discrete text entry technique. In this method, the user employs their right foot to move a cursor and select characters, while the left foot is responsible for confirming the selection with a toe tap. This typing interface features a single cursor, which can only be controlled by the right foot, as illustrated in Figure 7b. This design choice serves two main purposes.

First, akin to the principles governing the cooperation between adjacent motors in Guiard's kinematic chain model [17], *FeetAsymTap* allocates distinct tasks to the left and right feet, resembling the different positions of motors within a chain. This helps mitigate potential conflicts or confusion between selection and confirmation during text entry. Unlike hands, where users can perform pointing and selection almost simultaneously, the limited flexibility of feet requires a more sequential execution. Pointing, involving larger-scale movements, is assigned to the right foot (usually the dominant foot [14]), while the left foot confirms the selection. This division of labor enhances user operational efficiency. This division of labor, based on the right foot's suitability for finer actions, helps enhance user operational efficiency.

Second, the design considers the ease of performing toe raises when the leg is perpendicular to the ground in a seated posture. However, as the leg is moved backward in the seated position, performing toe taps becomes more challenging. This is primarily due to the relaxed or weaker state of the relevant muscle group and the shift in the body's center of gravity [50]. Considering the dominant foot typically exhibits greater effectiveness in psychomotor aspects compared to the non-dominant foot [41], we strategically assigned the right foot, the dominant foot for most individuals, to move on the floor.

5.3 FeetGestureTap

FeetGestureTap is an asymmetric bipedal text entry technique that combines swipe-based and tap-based methods. In this technique, the right foot is primarily responsible for word-level gestures, while the left foot handles character input with the tap-based approach. The user initiates a word-level gesture by performing a toe tap with their right foot, signaling the start of the gesture. The shape of the

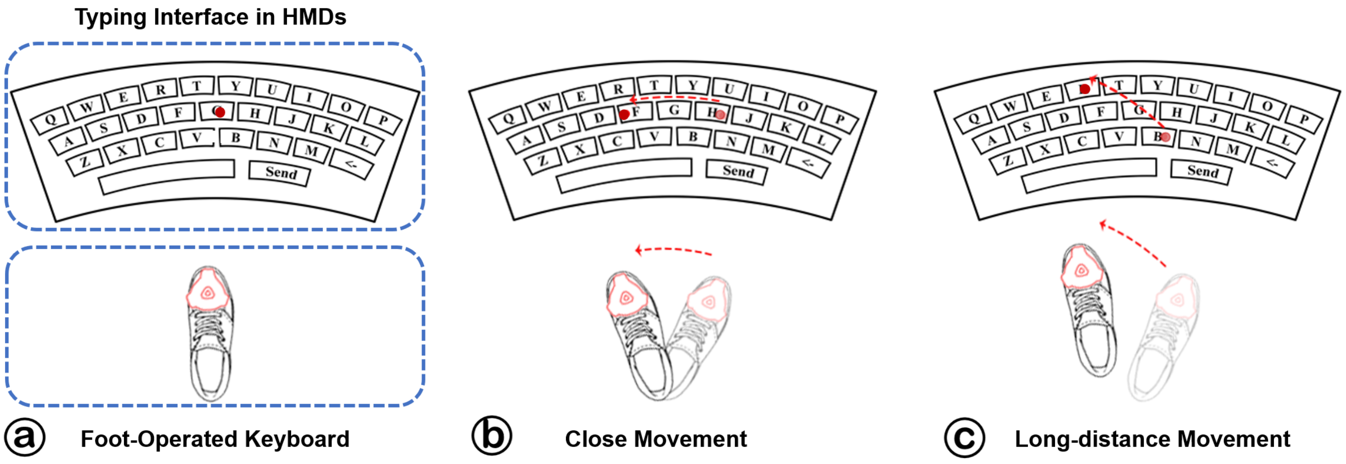

Figure 6: (a) The arched Qwerty keyboard is displayed in the HMD to provide visible feedback for the user. The foot-operated keyboard on the floor is also arch-shaped but is invisible. (b) Key 'F' is close to the location of the current cursor (key 'H'), so the user can simply pivot the ankle joint for a close movement. (c) The target key 'R' is far from the cursor (key 'B'), so the user needs to employ a combination of ankle joint rotation and a slight knee rotation to effectively reach the key 'R' to achieve long-distance movement.

word's gesture is then traced by sliding the right foot accordingly, and the gesture is completed with another tap, indicating the end of the gesture. Importantly, when the right foot draws a gesture, the cursor controlled by the left foot can move but cannot perform tap selections. This design ensures that both feet have distinct and complementary roles in the text entry process.

The foot-operated keyboard and typing interface in HMDs used for *FeetGestureTap* is a rectangle Qwerty keyboard. However, due to the potential for collision between the two feet when they are both in motion, special consideration is given to prevent such interference. For instance, if the right foot is placed on the key 'g', and the left foot attempts to tap on the letter 'i', the left foot may be blocked by the right foot on its way to the key 'i'. To address this issue, the positions of the foot-operated keyboards for the two feet are intentionally adjusted to be different, as depicted in Figure 7c. This asymmetrical arrangement helps avoid any potential interference or collision between the two feet as they move and interact with their respective keys.

To strike a balance between minimizing leg movements and preventing the 'fat finger' phenomenon [75] that can occur with excessively small keyboard sizes, each foot's foot-operated keyboard is set to a size of 60cm × 30cm, with individual letter keys measuring 5cm × 5cm (see Figure 7c). The virtual keyboard in the VR interface is placed 10m away from the center of the user's field of view, consistent with the keyboards used in the preliminary study.

6 STUDY 1: EVALUATION OF THE THREE USER-INSPIRED BIPED-BASED TEXT ENTRY TECHNIQUES

This study aims to evaluate the performance and user experience of the three biped-based text entry techniques.

6.1 Participants

Eighteen participants (8 males; 10 females) between the ages of 18-27 ($M = 22.89, SD = 2.42$) and heights of 158cm - 186cm ($M = 173.18, SD = 9.45$) were recruited from the same university campus to participate in this study.

6.2 Experiment Design and Procedure

We used a within-subjects design with TECHNIQUE as the independent variable. A Latin-Square approach was employed to counterbalance the sequence of the three techniques. For each technique, participants were tasked with transcribing a set of 12 sentences sourced from MacKenzie and Soukoreff's phrase set [35]. These sentences were selected at random, ensuring no repetitions. The initial two sentences were designated for training purposes and not recorded. Subsequently, the next ten sentences constituted formal trials, and their outcomes were recorded for subsequent analysis. After finishing typing for each technique, participants were required to fill out the NASA-TLX and PSSUQ questionnaires. Participants could take a break of at least three minutes between each technique or longer if requested. The entire experiment lasted approximately 40 minutes. As a result, a total of 540 trials were included in the analysis dataset (= 18 participants × 3 text entry techniques × 10 sentences).

6.3 Results

The results of Shapiro-Wilk tests indicated that TER, NCER, NASA-TLX data were not normally distributed ($p < .05$). Thus, we applied RM-ANOVAs for one-dimensional normally distributed data, MANOVA for multi-dimensional normally distributed data (NASA-TLX data), and Friedman tests for non-normally distributed data (TER and NCER). we reported the degrees of freedom with Greenhouse-Geisser correction (when $\epsilon < .75$) or Huynh-Feldt correction ($\epsilon > .75$). Effect sizes were reported using partial eta squared (η_p^2) for

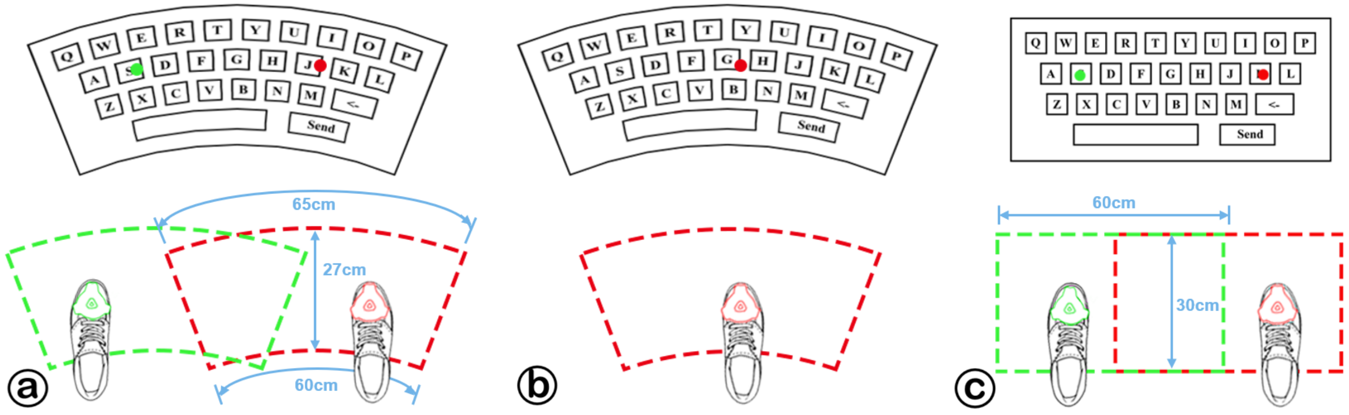

Figure 7: The typing interface in the VR HMDs and foot-operation keyboard of (a) *FeetSymTap*, the arched typing area for the left foot overlaps with that of the right foot by approximately 1/3. (b) *FeetAsymTap*, only one arched typing area beneath the right foot. (c) *FeetGestureTap*, there are rectangular typing areas beneath both the left and right feet, with approximately one-third overlap between the two. The actual location of the HTC VIVE Tracker is the size of the foot-operation keyboard. The valid active area of the HTC VIVE Tracker corresponds to the typing area that matches its color.

ANOVA tests and Kendall's W for Friedman tests. Post-hoc pairwise comparisons with Bonferroni correction were used if significant differences were identified.

6.3.1 Entry Rate and Error Rate. The RM-ANOVA yielded TECHNIQUE had a significant effect on entry rate ($F_{2,34} = 6.444, p = .004, \eta_p^2 = .275$), as shown in Figure 8a. Post-hoc pairwise comparisons indicated the entry rate of *FeetGestureTap* ($M = 9.16, SD = 2.16$) was significantly slower than *FeetSymTap* ($M = 11.12, SD = 1.94$) ($p = .006$) and *FeetAsymTap* ($M = 11.80, SD = 1.94$) ($p = .030$).

Friedman tests indicated that TECHNIQUE had a significant main effect on TER ($\chi_3^2 = 6.282, p = .043, W = .174$), but no significant effect on NCER ($p > .05$) (see Figure 8b and c, accordingly). Post-hoc tests found *FeetAsymTap* ($Mdn = 3.10\%$) led to lower error rates than *FeetSymTap* ($Mdn = 3.59\%$) ($p = .020$).

6.3.2 Usability. RM-ANOVA tests revealed significant differences in PSSUQ overall scores ($F_{1.262,21.454} = 6.454, p = .014, \eta_p^2 = .275$), system usefulness scores ($F_{2,34} = 5.127, p = .011, \eta_p^2 = .232$) and information quality ($F_{1.143,19.427} = 6.661, p = .0115, \eta_p^2 = .282$) among the three techniques, as shown in Figure 9a. Post-hoc pairwise comparisons revealed that the PSSUQ overall score ($M = 2.59, SD = 1.17$) and system usefulness ($M = 2.44, SD = 1.03$) of *FeetGestureTap* were higher than them of *FeetSymTap* ($M = 1.93, SD = 0.45$; $p = .041$ for overall scores and $M = 2.02, SD = 0.66$; $p = .029$ for system usefulness). The overall score of *FeetGestureTap* was significantly higher than that of *FeetAsymTap* ($M = 1.82, SD = 0.68$) ($p = .047$).

6.3.3 Perceived Workload. Figure 9b shows the NASA-TLX scores for the three biped-based text entry techniques. MANOVAs revealed a significant difference in perceived workload ($F = 29.034, p < .001, Wilks'\Lambda = .064, \eta_p^2 = .936$). For each dimension of NASA-TLX, RM-ANOVAs showed significant effects in physical demand ($F_{2,34} = 3.570, p = .039, \eta_p^2 = .174$), temporal demand ($F_{2,34} = 3.910, p = .030, \eta_p^2 = .187$), effort ($F_{2,34} = 5.989, p = .006, \eta_p^2 = .261$), performance ($F_{2,34} = 16.301, p < .001, \eta_p^2 = .490$), and frustration ($F_{2,34} = 12.840, p = .001, \eta_p^2 = .430$). Post-hoc tests indicated that *FeetAsymTap* ($M = 41.11, SD = 15.58$) required less physical demand than *FeetGestureTap* ($M = 52.78, SD = 19.50$) ($p = .044$). Further, participants were less satisfied with their performance on *FeetGestureTap* ($M = 40.00, SD = 15.62$) than *FeetAsymTap* ($M = 24.44, SD = 13.71$) and *FeetSymTap* ($M = 25.83, SD = 10.88$) (both $p < .001$), they felt more effort was put using *FeetGestureTap* ($M = 53.61, SD = 19.84$) than *FeetAsymTap* ($M = 39.44, SD = 12.94$) ($p = .031$) and *FeetSymTap* ($M = 42.22, SD = 11.91$) ($p = .032$), and suffered greater frustration with *FeetGestureTap* ($M = 37.78, SD = 20.95$) than *FeetAsymTap* ($M = 22.22, SD = 10.65$) ($p = .001$) and *FeetSymTap* ($M = 24.72, SD = 15.95$) ($p = .003$).

6.4 Discussion

The two tap-based biped text entry techniques demonstrated entry rates of 11.12 WPM and 10.80 WPM, respectively, as Figure 8 shows. These rates are comparable to other hands-free typing techniques, including dwell-based approaches (10.20 WPM [58]; 10.59 WPM [75]; 11.18 WPM [34]), BlinkType (13.47 WPM [34]), and NeckType (11.18 WPM [34]). Meanwhile, *FeetGestureTap* (9.16 WPM), the word-gesture-based method, exhibited a slightly slower entry rate than the two tap-based techniques. Overall, the error rates across the three techniques were low. The NCER of the three biped-based techniques was not significantly different. That means there is no difference between the three techniques regarding their impact on user-initiated error correction.

Users rated the usability of all three biped-based typing techniques highly, with the mean PSSUQ scores lower than 3 (see Figure 9a). All three techniques show acceptable workload levels (most scored below 50) (see Figure 9b) [13]. Across all three techniques, participants reported mental workloads that were consistently in the range of 30 points, indicating a moderate level of cognitive engagement during the typing tasks. The workloads associated with the *FeetAsymTap* and *FeetSymTap* generally fall below 40, except

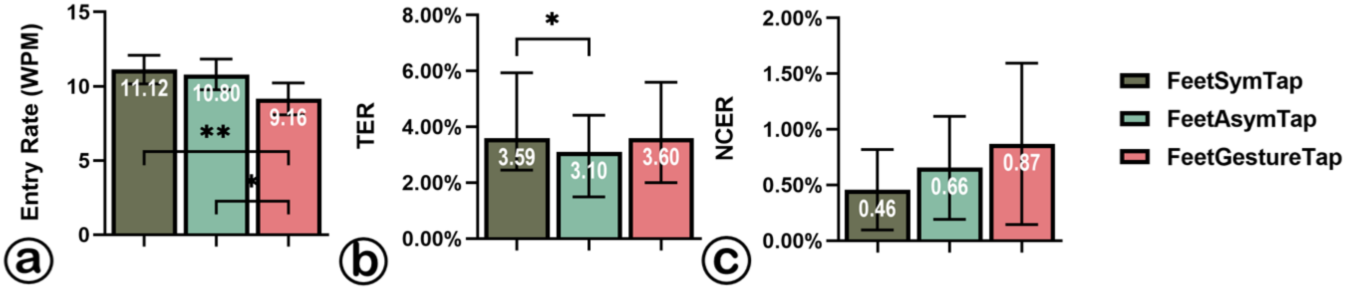

Figure 8: The means of (a) Entry rate, (b) TER, and (c) NCER. *, **, and * represent a .001, .01, and .05 significance level (Bonferroni-adjusted), respectively. The same marking scheme is used in Figure 9.**

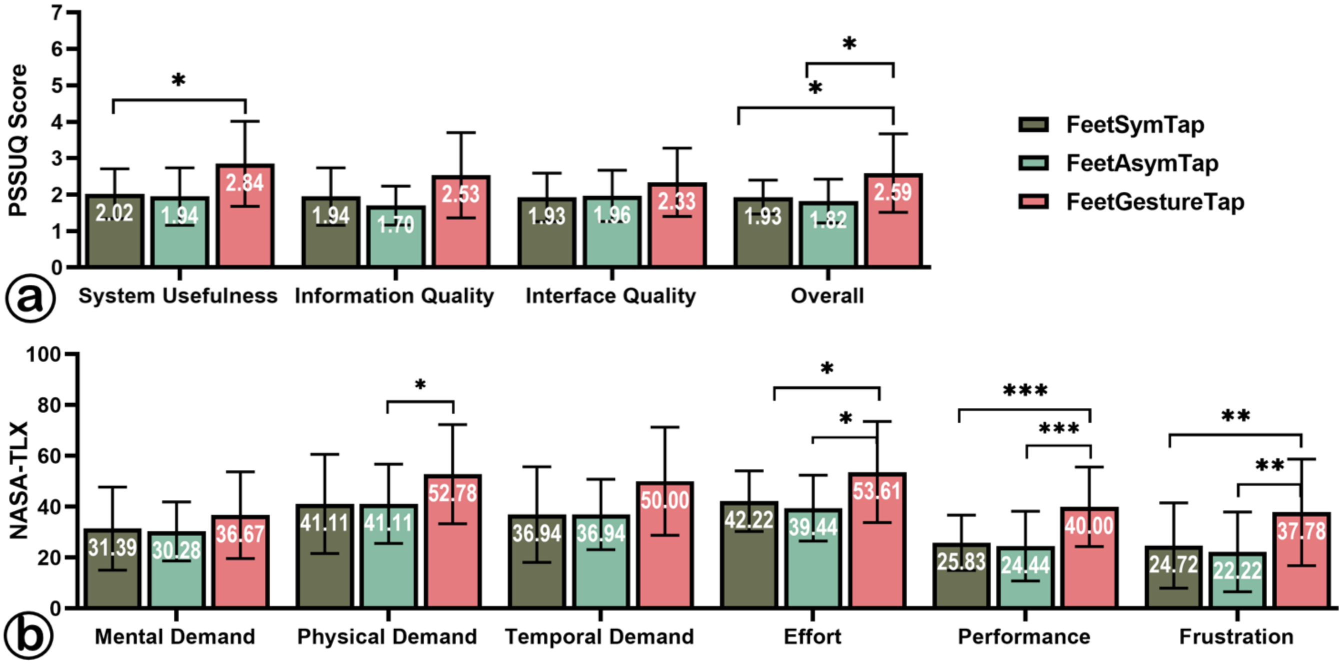

Figure 9: The means of (a) PSSUQ Score and (b) NASA-TLX scores.

for a slightly higher score in physical demands, just above 40 (Any workload exceeding 40 is deemed high [25]). In contrast, the NASA-TLX score for *FeetGestureTap* tends to be slightly higher. In previous VR text entry research (e.g. [16, 26, 43, 58]), the NASA-TLX scores were generally above 40. The NASA-TLX scores of *FeetAsymTap* and *FeetSymTap* are comparable to the NASA-TLX scores of these text entry techniques. This suggests that these two foot-based text entry techniques might also offer efficiency or ease of learning and mastery during user interaction and use. The slightly higher score for *FeetGestureTap* may imply increased demands in certain aspects, possibly involving more physical movement or cognitive burden. The consistency in the NASA-TLX scores suggests that typing in a VR environment is a challenging task, potentially due to the complexity of VR environments, interaction dynamics, and inherent challenges associated with text input. The slightly higher score for *FeetGestureTap* may imply increased demands in certain aspects, possibly involving more physical movement or cognitive burden.

The outcomes about text entry performance and subjective feedback affirm the proficiency of the three biped-based techniques in achieving an acceptable pace of input for hands-free text entry within the VR context (**RQ2**).

6.4.1 Symmetric vs. Asymmetric Bipedal Tap-based Techniques. The tap-based text entry techniques, *FeetAsymTap* and *FeetSymTap*, involve similar foot actions, showcasing simplicity in their execution, whether the technique is asymmetrical or symmetrical. In tasks requiring basic actions or movements that mirror each other, such as stepping or rotating the toes of both feet, the feet generally demonstrate a similar level of proficiency [14]. Thus, users did not perceive significant differences in typing rates or task difficulty between the two techniques.

The total error rate (TER) of *FeetAsymTap* (3.10%) was marginally lower compared to *FeetSymTap* (3.59%). This distinction can be attributed to the higher level of coordination required for tapping and positioning on a single foot, and the non-dominant foot has less mobility than the dominant one [41]. Precise control over foot movement is essential to position and tap accurately. In *FeetAsymTap*, the right foot, serving as the dominant foot for users, performs the finer pointing action, which helps reduce the occurrence of errors. In *FeetSymTap*, without specific functionalities assigned to each foot, there might be increased confusion between different actions, subsequently affecting accuracy. This issue arises due to the constant need to transition between toe-tapping and lateral floor movement, introducing added complexity that may contribute to errors.

It is essential to acknowledge that people's feet exhibit less flexibility and coordination compared to their hands. In the context of hand-based text entry, pointing and selecting actions can be considered almost equal in their significance and often occur concurrently. However, when translating the text entry task to the feet, the task takes place sequentially.

The text entry task delineates pointing and tapping as two hierarchical actions. In *FeetAsymTap*, the tapping of the left foot action follows the pointing of the right foot, seemingly deviating from the left hand (non-dominant foot) precedence principle in the Guiard's kinematic chain model of asymmetrical bimanual tasks [17]. This divergence occurs because the pointing action of the lower limb, compared to the toe-tapping action, is more intricate and holds a higher significance in typing tasks. Due to greater precision/control in the right foot (dominant foot), the right foot undertakes the more complex task.

*6.4.2 Tapping vs. Gesturing for Foot-based Text Entry (**RQ3**).* The comparison between the swipe-based technique (*FeetGestureTap*) and the two tap-based techniques (*FeetSymTap* and *FeetAsymTap*) reveals differences in entry rates (see Figure 8a) and user perceptions (see Figure 9a and b). *FeetGestureTap* showed a slower entry rate than the tap-based techniques, and participants found it less usable and useful, according to PSSUQ results. NASA-TLX scores indicated higher physical demand, effort, performance, and frustration with *FeetGestureTap*, suggesting increased participant effort without commensurate performance improvement.

In our preliminary study, we found no significant differences in text entry rates between standing and sitting postures for both tap and swipe mechanisms (Section 4.5). This lack of disparity can be attributed to the continuous single-leg input required in the tapping mechanism, involving frequent transitions between pointing and selection actions, whereas gesturing, despite its continuous nature, involves a lower frequency of switching between pointing and selection actions. However, when both feet are available for alternating use in a sitting posture, the tapping mechanism for text entry demonstrates superior performance. This suggests that the efficiency of text input using tapping versus gesturing mechanisms is influenced by the alternating use of both feet. *FeetSymTap* allows users to freely decide which foot to use, and the user is free to decide which foot to use because both feet have the same function. With *FeetAsymTap*, the user is forced to alternate the use of both feet; that is, the left foot needs to be tapped after each pointing of the right foot. The feet are mainly used in an alternating form in daily life, such as walking and climbing stairs. This means that alternate use is more in alignment with daily behavior.

6.4.3 Ergonomic Keyboard Layout. The reduced physical demands with *FeetSymTap*, which can be attributed to its arched Qwerty keyboard, aligning with the foot's natural movement path. This ergonomic design minimizes the need for extended leg movements, enhancing ease and accuracy in tapping keys. Alternative keyboard layouts, like RingText [72] and PizzaText [76] that are based on a circular layout and Flower Text Entry [29] that follows a flower shaped layout, have further emphasized the significance of ergonomic design, showcasing improved text entry performance through layouts aligned with natural body movements. These approaches highlight the importance of ergonomic considerations in keyboard design, promoting user comfort and ease of use, potentially enhancing typing speeds and accuracy.

6.4.4 Non-overlapping Input Spaces for Indirect Interaction. With *FeetSymTap*, the intentionally misaligned design of the foot-operated keyboards, aimed at preventing potential obstructions between the movement of both feet, did not result in reported typing difficulties. Participants interacting with the indirect interface prioritized observing on-screen cursors within their HMD to interpret characters for selection, similar to the mechanism observed in mouse operations. This implicit reliance on cursor positions on the VR display suggests that users may not heavily consider the precise positioning of their feet when operating the foot-operated keyboards. This alignment discrepancy between the virtual display and the actual placement of the keyboards did not cause operational discomfort or cognitive challenges for participants, which aligns with Seinfeld et al.'s [56] findings that during indirect interactions using a physical stick to reach a distant object or a virtual cursor to select an icon, the user can transfer motor control of a specific body part to a mechanical or virtual end-effector without a significant impact on performance.

6.5 Lessons Learned from the User Study

- **L5.** The outcomes about text entry performance and subjective feedback affirm the proficiency of the three biped-based techniques in achieving an acceptable pace of input for hands-free text entry within the VR context (**RQ2**).
- **L6.** Coordinating tapping and positioning on the same foot is demanding and can be disruptive to smooth movements, potentially affecting accuracy. The frequent transition between toe-tapping and lateral movement introduces complexity, which may lead to errors.
- **L7.** The bipedal tap-based techniques significantly outperformed the swipe-based technique in terms of typing entry rate and subjective feedback mainly due to the alternative use of the feet. As such, alternate foot use is an effective way to mitigate foot fatigue.

6.6 Potential (Un)Suitable Application Scenarios of Foot-based Text Entry

The performance of the three biped-based text input techniques demonstrates the feasibility of doing lightweight input tasks in a seated position. Foot-based typing could address issues related to head motion-induced discomfort, free upper limbs to allow a more relaxed interaction (e.g., when watching a movie in a seated position), and is inconspicuous, which is especially useful in public scenarios.

Head movements might cause issues such as motion sickness [32, 37, 38] or simulator sickness [54, 63, 67], particularly in mobile environments such as being in a subway and bus. When both hands are occupied in these mobile environments, foot-based text entry could be considered a possible and viable input method (see Figure 2a for an example).

Foot typing in some seated positions, such as when watching movies, mitigates the disruption of comfort and immersion by utilizing the natural position of users' feet on the floor, allowing for a relaxed and freely movable posture without involving the upper limbs. Moreover, in some environments, such as a desk/table in front of a user or dim lights, foot-based typing is unobtrusive, making it more socially acceptable in public areas [62].

In addition, foot typing introduces a viable approach to deskless interactions. In an unrestricted office environment without the constraints of a traditional desk setup, users have the freedom to choose their seating arrangements, such as opting for a comfortable sofa or couch to sit on. Without a traditional office desk, typing in mid-air can lead to hand fatigue [22, 68], making foot typing an effective supplement or alternative. Users have the flexibility to choose to use foot typing exclusively or switch to hand typing when fatigue sets in (see Figure 2b).

In daily life, we commonly adopt three postures: sitting, standing, and walking. Foot-based text entry is most suitable for users in a seated position, where the feet are relaxed. While standing is also feasible, it may not be ideal for prolonged use due to balance considerations and limited foot movement. Walking, with its focus on maintaining balance and awareness of surroundings, poses challenges for foot-operated text input, making it less optimal, especially in situations requiring heightened concentration.

7 DESIGN RECOMMENDATIONS FOR FOOT-BASED TEXT ENTRY (RQ2)

From the above results, we can distill the following five design considerations.

- **Minimizing leg movement**. Minimizing the extent of leg movement is crucial, as excessive leg movements can easily lead to fatigue, subsequently impacting typing performance. The performance and workload involved were optimized when the user used both feet and had a reduced range of foot motion, as observed in Study One.
- **Redesigning the keyboard layout**. Aligning the keyboard layout with the inherent motion paths of the legs and feet ensures that users can interact with the keyboard comfortably and efficiently, reducing the risk of strain or discomfort associated with unnatural movements. As pointed out in Section 6.4.3, the arched Qwerty keyboard is one of the reasons the two tap-based techniques (*FeetSymTap* and *FeetAsymTap*) led to lower physical demands due to the more natural accessibility of the target keys.
- **Alternating foot usage**. Encouraging users to alternate between feet for input can distribute the load more evenly and help reduce the risk of fatigue in a single leg. The typing performance greatly improved when using the three biped-based text entry feet, as the results of Study 1 showed.
- **Mapping tasks based on foot's dexterity**. The non-dominant foot is better suited for relatively simple tasks like tapping rather than complex tasks such as precise key positioning or tracing. The dominant foot can be assigned more complex tasks and those requiring precision. In the preliminary study, we found that users used the right foot (their dominant foot) more frequently in both standing and sitting positions, and fewer errors occurred when the purpose of the left and right feet was separated clearly (lower TER for *FeetAsymTap* compared to *FeetSymTap*).
- **Prioritizing simple actions**. Swipe selection is more complex and necessitates continuous and coordinated leg movements, which can be more physically demanding than tapping and is less suitable where there is a need to switch between two feet frequently. Users' feet are not as dexterous as their hands. As such, for repetitive tasks like text entry, simple, easy-to-perform actions look more suitable and preferred by users. That is one reason *FeetGestureTap* did not perform as well as *FeetSymTap* and *FeetAsymTap* in entry rate and subjective feedback.

8 LIMITATIONS AND FUTURE WORK

There are some limitations in our work that can help frame future lines of research.

First, We tracked users' foot movement and gestures using HTC VIVE Trackers, which is effective in controlled environments. However, using these trackers in public or mobile settings has limitations. Recent studies propose integrating smaller sensors onto shoes, utilizing technologies like pressure sensors and accelerometers for richer foot motion information [61]. Future research could explore diverse sensors to capture a broader range of foot movements, enabling the design of alternative text entry techniques for various scenarios.

Second, our work on foot-based typing in VR environments did not focus too much on the possibility of users accidentally triggering keyboard operations, given that this was not an issue in our first preliminary study. We made a basic distinction between the slight foot-lifting and toe-tapping actions. As this research represents an initial exploration into using feet for typing within VR settings, the risks associated with the potential accidental triggering of operations have not been examined in detail. Future work can delve deeper into understanding and mitigating the risk of inadvertently triggering actions during text entry via users' feet.

Third, the proposed foot-based techniques require accurate recognition of toe taps. However, we did not account for variations in user anatomical structure, types of footwear, and other factors that might impact the accuracy of recognizing toe taps. While these aspects did not affect the techniques' performance, there could be instances where they could have an effect. As such, future studies can consider different recognition methods to minimize issues arising from large variations in users' body types and footwear.

Fourth, our study is focused on VR, and while the findings can potentially be extended to MR and AR HMDs, we did not conduct experiments in either AR or MR environments. Future research could factor in elements specific to AR and MR use scenarios. Such exploration will enable a more comprehensive understanding of foot-based text entry across the different spectrums of extended reality systems, including AR and MR, and design techniques that are efficient, context-fit, and usable.

9 CONCLUSION

Our work explored the challenges and possibilities of text entry in Virtual Reality (VR) environments, particularly when users' hands are occupied or in situations where they are unsuitable to be used. These contexts necessitate alternative input methods to make text entry possible and enhance users' experience. Our work investigated foot-based input as an alternative. We first conducted a preliminary study to investigate the feasibility of foot-based text entry and a second study to evaluate the performance of three proposed foot-based techniques (*FeetSymTap*, *FeetAsymTap*, and *FeetGestureTap*) that included (1) standing and seated postures, as well as (2) tap and swipe mechanisms. Results from the user studies show that foot-based input is feasible and practical for text entry in VR scenarios. In addition, they indicate that tap-based techniques (*FeetSymTap* and *FeetAsymTap*) allowed faster typing with lower workloads than *FeetGestureTap*, a gesture-based technique. Similarly, *FeetAsymTap* achieved a slightly lower total error rate compared with *FeetSymTap*. Our investigation shed light on the potential of foot-based text entry techniques in VR and highlighted the nuances of their performance. These insights can serve as a foundation for future research to refine the design of foot-based input, explore further optimization strategies, and develop personalized solutions that cater to individual user preferences and characteristics. As VR continues to evolve and expand its applications and target users, our findings contribute to the ongoing work on enhancing user experiences with VR systems and have the potential to be extended to AR/MR scenarios.

ACKNOWLEDGMENTS

The authors thank the participants who volunteered their time to join the user studies. We also thank the reviewers whose insightful comments and suggestions helped improve our paper. This work was partially funded by the Suzhou Municipal Key Laboratory for Intelligent Virtual Engineering (#SZS2022004) and the National Natural Science Foundation of China (#62272396).

REFERENCES

[1] 2011. The Open American National Corpus. https://anc.org.

[2] Jiban Adhikary and Keith Vertanen. 2021. Typing on Midair Virtual Keyboards: Exploring Visual Designs and Interaction Styles. In *Human-Computer Interaction – INTERACT 2021*, Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, and Kori Inkpen (Eds.). Springer International Publishing, Cham, 132–151.

[3] Garrett Benoit, G. Michael Poor, and Alvin Jude. 2017. Bimanual Word Gesture Keyboards for Mid-Air Gestures. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI EA '17)*. Association for Computing Machinery, New York, NY, USA, 1500–1507. https://doi.org/10.1145/3027063.3053137

[4] Xiaojun Bi, Ciprian Chelba, Tom Ouyang, Kurt Partridge, and Shumin Zhai. 2012. Bimanual Gesture Keyboard. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology* (Cambridge, Massachusetts, USA) *(UIST '12)*. Association for Computing Machinery, New York, NY, USA, 137–146. https://doi.org/10.1145/2380116.2380136

[5] Costas Boletsis and Stian Kongsvik. 2019. Controller-based Text-input Techniques for Virtual Reality: An Empirical Comparison. *International Journal of Virtual Reality* 19, 3 (Oct. 2019), 2–15. https://doi.org/10.20870/IJVR.2019.19.3.2917

[6] Sibo Chen, Junce Wang, Santiago Guerra, Neha Mittal, and Soravis Prakkamakul. 2019. Exploring Word-Gesture Text Entry Techniques in Virtual Reality. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3290607.3312762

[7] Justin Cuaresma and I Scott MacKenzie. 2013. A study of variations of Qwerty soft keyboards for mobile phones. In *Proceedings of the International Conference on Multimedia and Human-Computer Interaction-MHCI*. International ASET, Ottawa, Canada, 126.1–126.8.

[8] Wenzhe Cui, Jingjie Zheng, Blaine Lewis, Daniel Vogel, and Xiaojun Bi. 2019. HotStrokes: Word-Gesture Shortcuts on a Trackpad. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300395

[9] Wenzhe Cui, Suwen Zhu, Zhi Li, Zheer Xu, Xing-Dong Yang, IV Ramakrishnan, and Xiaojun Bi. 2021. BackSwipe: Back-of-Device Word-Gesture Interaction on Smartphones. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 196, 12 pages. https://doi.org/10.1145/3411764.3445081

[10] Krzysztof Dobosz and Mirosław Trzcionkowski. 2020. Text Input with Foot Gestures Using the Myo Armband. In *Computers Helping People with Special Needs*, Klaus Miesenberger, Roberto Manduchi, Mario Covarrubias Rodriguez, and Petr Peňáz (Eds.). Springer International Publishing, Cham, 335–342.

[11] Tafadzwa Joseph Dube, Kevin Johnson, and Ahmed Sabbir Arif. 2022. Shapeshifter: Gesture Typing in Virtual Reality with a Force-Based Digital Thimble. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 230, 9 pages. https://doi.org/10.1145/3491101.3519679

[12] John Dudley, Hrvoje Benko, Daniel Wigdor, and Per Ola Kristensson. 2019. Performance Envelopes of Virtual Keyboard Text Input Strategies in Virtual Reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. Beijing, China, 289–300. https://doi.org/10.1109/ISMAR.2019.00027

[13] Maren Helene Eitrheim and Alexandra Fernandes. 2016. The NASA Task Load Index for rating workload acceptability. In *Human Factors and User Needs in Transport, Control, and the Workplace-HFES-Europe Annual Meeting*. Prague. https://doi.org/10.13140/RG.2.2.22989.64484

[14] Yasmin Felberbaum and Joel Lanir. 2018. Better Understanding of Foot Gestures: An Elicitation Study. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173908

[15] Joshua Goodman, Gina Venolia, Keith Steury, and Chauncey Parker. 2002. Language Modeling for Soft Keyboards. In *Proceedings of the 7th International Conference on Intelligent User Interfaces* (San Francisco, California, USA) *(IUI '02)*. Association for Computing Machinery, New York, NY, USA, 194–195. https://doi.org/10.1145/502716.502753

[16] Jens Grubert, Lukas Witzani, Eyal Ofek, Michel Pahud, Matthias Kranz, and Per Ola Kristensson. 2018. 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)*. Tuebingen/Reutlingen, Germany, 159–166. https://doi.org/10.1109/VR.2018.8446059

[17] Yves Guiard. 1987. Asymmetric division of labor in human skilled bimanual action: the kinematic chain as a model. *Journal of motor behavior* 19:4 (1987), 486–517. https://doi.org/10.1080/00222895.1987.10735426

[18] Aakar Gupta, Cheng Ji, Hui-Shyong Yeo, Aaron Quigley, and Daniel Vogel. 2019. RotoSwype: Word-Gesture Typing Using a Ring. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300244

[19] John Paulin Hansen, Vijay Rajanna, I. Scott MacKenzie, and Per Bækgaard. 2018. A Fitts' Law Study of Click and Dwell Interaction by Gaze, Head and Mouse with a Head-Mounted Display. In *Proceedings of the Workshop on Communication by Gaze Interaction* (Warsaw, Poland) *(COGAIN '18)*. Association for Computing Machinery, New York, NY, USA, Article 7, 5 pages. https://doi.org/10.1145/3206343.3206344

[20] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. https://doi.org/10.1177/154193120605000909

[21] Jay Henderson, Jessy Ceha, and Edward Lank. 2020. *STAT: Subtle Typing Around the Thigh for Head-Mounted Displays*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3379503.3403549

[22] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-Air Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1063–1072. https://doi.org/10.1145/2556288.2557130

[23] Florian Kern, Florian Niebling, and Marc Erich Latoschik. 2023. Text Input for Non-Stationary XR Workspaces: Investigating Tap and Word-Gesture Keyboards in Virtual and Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 29, 5 (2023), 2658–2669. https://doi.org/10.1109/TVCG.2023.3247098

[24] Konstantin Klamka, Andreas Siegel, Stefan Vogt, Fabian Göbel, Sophie Stellmach, and Raimund Dachselt. 2015. Look & Pedal: Hands-Free Navigation in Zoomable Information Spaces through Gaze-Supported Foot Input. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (Seattle, Washington, USA) *(ICMI '15)*. Association for Computing Machinery, New York, NY, USA, 123–130. https://doi.org/10.1145/2818346.2820751

[25] Beverly G. Knapp and Mary Jane Hall. 1990. Human Performance Concerns for the TRACKWOLF System. https://api.semanticscholar.org/CorpusID:107379503

[26] Pascal Knierim, Thomas Kosch, Johannes Groschopp, and Albrecht Schmidt. 2020. Opportunities and Challenges of Text Input in Portable Virtual Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382920

[27] Per-Ola Kristensson and Shumin Zhai. 2004. SHARK2: A Large Vocabulary Shorthand Writing System for Pen-Based Computers. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology* (Santa Fe, NM, USA) *(UIST '04)*. Association for Computing Machinery, New York, NY, USA, 43–52. https://doi.org/10.1145/1029632.1029640

[28] Chandan Kumar, Ramin Hedeshy, I. Scott MacKenzie, and Steffen Staab. 2020. TAGSwipe: Touch Assisted Gaze Swipe for Text Entry. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376317

[29] Jiaye Leng, Lili Wang, Xiaolong Liu, Xuehuai Shi, and Miao Wang. 2022. Efficient Flower Text Entry in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* 28, 11 (2022), 3662–3672. https://doi.org/10.1109/TVCG.2022.3203101

[30] James R. Lewis. 1991. Psychometric Evaluation of an After-Scenario Questionnaire for Computer Usability Studies: The ASQ. *SIGCHI Bull.* 23, 1 (jan 1991), 78–81. https://doi.org/10.1145/122672.122692

[31] James R. Lewis. 1992. Psychometric Evaluation of the Post-Study System Usability Questionnaire: The PSSUQ. *Proceedings of the Human Factors Society Annual Meeting* 36, 16 (1992), 1259–1260. https://doi.org/10.1177/154193129203601617

[32] Jingyi Li, Agnes Reda, and Andreas Butz. 2021. Queasy Rider: How Head Movements Influence Motion Sickness in Passenger Use of Head-Mounted Displays. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Leeds, United Kingdom) *(AutomotiveUI '21)*. Association for Computing Machinery, New York, NY, USA, 28–38. https://doi.org/10.1145/3409118.3475137

[33] Xueshi Lu, Difeng Yu, Hai-Ning Liang, and Jorge Goncalves. 2021. IText: Hands-Free Text Entry on an Imaginary Keyboard for Augmented Reality Systems. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21)*. Association for Computing Machinery, New York, NY, USA, 815–825. https://doi.org/10.1145/3472749.3474788

[34] Xueshi Lu, Difeng Yu, Hai-Ning Liang, Wenge Xu, Yuzheng Chen, Xiang Li, and Khalad Hasan. 2020. Exploration of Hands-free Text Entry Techniques For Virtual Reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. Porto de Galinhas, Brazil, 344–349. https://doi.org/10.1109/ISMAR50242.2020.00061

[35] I. Scott MacKenzie and R. William Soukoreff. 2003. Phrase Sets for Evaluating Text Entry Techniques. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI EA '03)*. Association for Computing Machinery, New York, NY, USA, 754–755. https://doi.org/10.1145/765891.765971

[36] Anders Markussen, Mikkel Rønne Jakobsen, and Kasper Hornbæk. 2014. Vulture: A Mid-Air Word-Gesture Keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1073–1082. https://doi.org/10.1145/2556288.2556964

[37] Omar Merhi, Elise Faugloire, Moira Flanagan, and Thomas A. Stoffregen. 2007. Motion Sickness, Console Video Games, and Head-Mounted Displays. *Human Factors* 49, 5 (2007), 920–934. https://doi.org/10.1518/001872007X230262

[38] Diego Monteiro, Hai-Ning Liang, Xiaohang Tang, and Pourang Irani. 2021. Using Trajectory Compression Rate to Predict Changes in Cybersickness in Virtual Reality Games. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 138–146. https://doi.org/10.1109/ISMAR52148.2021.00028

[39] Florian Müller, Joshua McManus, Sebastian Günther, Martin Schmitz, Max Mühlhäuser, and Markus Funk. 2019. Mind the Tap: Assessing Foot-Taps for Interacting with Head-Mounted Displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300707

[40] Taihei Ogitani, Yoshitaka Arahori, Yusuke Shinyama, and Katsuhiko Gondow. 2018. 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)*. 342–349. https://doi.org/10.1109/AINA.2018.00059

[41] Toni Pakkanen and Roope Raisamo. 2004. Appropriateness of Foot Interaction for Non-Accurate Spatial Tasks. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria) *(CHI EA '04)*. Association for Computing Machinery, New York, NY, USA, 1123–1126. https://doi.org/10.1145/985921.986004

[42] Diogo Pedrosa and Maria da Graça C. Pimentel. 2014. Text Entry Using a Foot for Severely Motor-Impaired Individuals. In *Proceedings of the 29th Annual ACM Symposium on Applied Computing* (Gyeongju, Republic of Korea) *(SAC '14)*. Association for Computing Machinery, New York, NY, USA, 957–963. https://doi.org/10.1145/2554850.2554948

[43] Duc-Minh Pham and Wolfgang Stuerzlinger. 2019. HawKEY: Efficient and Versatile Text Entry for Virtual Reality. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 21, 11 pages. https://doi.org/10.1145/3359996.3364265

[44] Yuan Yuan Qian and Robert J. Teather. 2017. The Eyes Don't Have It: An Empirical Comparison of Head-Based and Eye-Based Selection in Virtual Reality. In *Proceedings of the 5th Symposium on Spatial User Interaction* (Brighton, United Kingdom) *(SUI '17)*. Association for Computing Machinery, New York, NY, USA, 91–98. https://doi.org/10.1145/3131277.3132182

[45] Kari-Jouko Räihä and Saila Ovaska. 2012. An Exploratory Study of Eye Typing Fundamentals: Dwell Time, Text Entry Rate, Errors, and Workload. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 3001–3010. https://doi.org/10.1145/2207676.2208711

[46] Vijay Rajanna and Tracy Hammond. 2016. GAWSCHI: Gaze-Augmented, Wearable-Supplemented Computer-Human Interaction. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications* (Charleston, South Carolina) *(ETRA '16)*. Association for Computing Machinery, New York, NY, USA, 233–236. https://doi.org/10.1145/2857491.2857499

[47] Vijay Rajanna, Murat Russel, Jeffrey Zhao, and Tracy Hammond. 2022. PressTapFlick: Exploring a gaze and foot-based multimodal approach to gaze typing. *International Journal of Human-Computer Studies* 161 (2022), 102787. https://doi.org/10.1016/j.ijhcs.2022.102787

[48] Emily Rickel, Kelly Harris, Erika Mandile, Anthony Pagliari, Jessyca L. Derby, and Barbara S. Chaparro. 2022. Typing in Mid Air: Assessing One- and Two-Handed Text Input Methods of the Microsoft HoloLens 2. In *Virtual, Augmented and Mixed Reality: Design and Development: 14th International Conference, VAMR 2022, Held as Part of the 24th HCI International Conference, HCII 2022, Virtual Event, June 26 – July 1, 2022, Proceedings, Part I*. Springer-Verlag, Berlin, Heidelberg, 357–368. https://doi.org/10.1007/978-3-031-05939-1_24

[49] Asbjørn Roaas and Gunnar B. J. Andersson. 1982. Normal Range of Motion of the Hip, Knee and Ankle Joints in Male Subjects, 30–40 Years of Age. *Acta Orthopaedica Scandinavica* 53, 2 (1982), 205–208. https://doi.org/10.3109/17453678208992202

[50] M.E. Roebroeck, C.A.M. Doorenbosch, J. Harlaar, R. Jacobs, and G.J. Lankhorst. 1994. Biomechanics and muscular activity during sit-to-stand transfer. *Clinical Biomechanics* 9, 4 (1994), 235–244. https://doi.org/10.1016/0268-0033(94)90004-3

[51] William Saunders and Daniel Vogel. 2015. The performance of indirect foot pointing using discrete taps and kicks while standing. In *Proceedings of the 41st Graphics Interface Conference* (Halifax, Nova Scotia, Canada) *(GI '15)*. Canadian Information Processing Society, CAN, 265–272.

[52] William Saunders and Daniel Vogel. 2016. Tap-Kick-Click: Foot Interaction for a Standing Desk. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems* (Brisbane, QLD, Australia) *(DIS '16)*. Association for Computing Machinery, New York, NY, USA, 323–333. https://doi.org/10.1145/2901790.2901815

[53] Yuan-Fu Shao, Masatoshi Chang-Ogimoto, Reinhard Pointner, Yu-Chih Lin, Chen-Ting Wu, and Mike Chen. 2016. SwipeKey: A Swipe-Based Keyboard Design for Smartwatches. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Florence, Italy) *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 60–71. https://doi.org/10.1145/2935334.2935336

[54] Rongkai Shi, Hai-Ning Liang, Yu Wu, Difeng Yu, and Wenge Xu. 2021. Virtual Reality Sickness Mitigation Methods: A Comparative Study in a Racing Game. *Proc. ACM Comput. Graph. Interact. Tech.* 4, 1, Article 8 (apr 2021), 16 pages. https://doi.org/10.1145/3451255

[55] Adalberto L. Simeone, Eduardo Velloso, Jason Alexander, and Hans Gellersen. 2014. Feet movement in desktop 3D interaction. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*. 71–74. https://doi.org/10.1109/3DUI.2014.6798845

[56] Antonella Maselli Sofia Seinfeld, Tiare Feuchtner and Jörg Müller. 2021. User Representations in Human-Computer Interaction. *Human–Computer Interaction*

36, 5-6 (2021), 400–438. https://doi.org/10.1080/07370024.2020.1724790

[57] R. William Soukoreff and I. Scott MacKenzie. 2003. 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* (Ft. Lauderdale, Florida, USA) *(CHI '03)*. Association for Computing Machinery, New York, NY, USA, 113–120. https://doi.org/10.1145/642611.642632

[58] Marco Speicher, Anna Maria Feit, Pascal Ziegler, and Antonio Krüger. 2018. Selection-Based Text Entry in Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3174221

[59] Yanbo Tao, Tin Lun Lam, Huihuan Qian, and Yangsheng Xu. 2012. A real-time intelligent shoe-keyboard for computer input. In *2012 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. 1488–1493. https://doi.org/10.1109/ROBIO.2012.6491179

[60] Eduardo Velloso, Jason Alexander, Andreas Bulling, and Hans Gellersen. 2022. Interactions Under the Desk: A Characterisation of Foot Movements for Input in a Seated Position. In *Human-Computer Interaction – INTERACT 2015*. Springer-Verlag, Berlin, Heidelberg, 384–401. https://doi.org/10.1007/978-3-319-22701-6_29

[61] Eduardo Velloso, Dominik Schmidt, Jason Alexander, Hans Gellersen, and Andreas Bulling. 2015. The Feet in Human–Computer Interaction: A Survey of Foot-Based Interaction. *ACM Comput. Surv.* 48, 2, Article 21 (sep 2015), 35 pages. https://doi.org/10.1145/2816455

[62] Maurizio Vergari, Tanja Kojić, Francesco Vona, Franca Garzotto, Sebastian Möller, and Jan-Niklas Voigt-Antons. 2021. Influence of Interactivity and Social Environments on User Experience and Social Acceptability in Virtual Reality. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. 695–704. https://doi.org/10.1109/VR50410.2021.00096

[63] Alexander D Walker, Eric R Muth, Fred S Switzer, and Adam Hoover. 2010. Head movements and simulator sickness generated by a virtual environment. *Aviation, space, and environmental medicine* 81, 10 (2010), 929–934. https://doi.org/10.3357/ASEM.2735.2010

[64] Tingjie Wan, Rongkai Shi, Wenge Xu, Yue Li, Katie Atkinson, Lingyun Yu, and Hai-Ning Liang. 2024. Hands-free multi-type character text entry in virtual reality. *Virtual Reality* 28, 1 (2024), 8. https://doi.org/10.1007/s10055-023-00902-z

[65] Tingjie Wan, Yushi Wei, Rongkai Shi, Junxiao Shen, Per Ola Kristensson, Katie Atkinson, and Hai-Ning Liang. 2024. 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* (2024), 1–11. https://doi.org/10.1109/TVCG.2024.3349428

[66] Cheng-Yao Wang, Wei-Chen Chu, Po-Tsung Chiu, Min-Chieh Hsiu, Yih-Harn Chiang, and Mike Y. Chen. 2015. PalmType: Using Palms as Keyboards for Smart Glasses. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Copenhagen, Denmark) *(MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 153–160. https://doi.org/10.1145/2785830.2785886

[67] Jialin Wang, Hai-Ning Liang, Diego Monteiro, Wenge Xu, and Jimin Xiao. 2023. Real-Time Prediction of Simulator Sickness in Virtual Reality Games. *IEEE Transactions on Games* 15, 2 (2023), 252–261. https://doi.org/10.1109/TG.2022.3178539

[68] Yahui Wang, Yueyang Wang, Jingzhou Chen, Yincheng Wang, Jie Yang, Ting Jiang, and Jibo He. 2021. Investigating the Performance of Gesture-Based Input for Mid-Air Text Entry in a Virtual Environment: A Comparison of Hand-Up versus Hand-Down Postures. *Sensors* 21, 5 (2021). https://doi.org/10.3390/s21051582

[69] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. *The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures.* Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[70] Wenge Xu, Hai-Ning Liang, Anqi He, and Zifan Wang. 2019. Pointing and Selection Methods for Text Entry in Augmented Reality Head Mounted Displays. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 279–288. https://doi.org/10.1109/ISMAR.2019.00026

[71] Wenge Xu, Hai-Ning Liang, Yuxuan Zhao, Difeng Yu, and Diego Monteiro. 2019. DMove: Directional Motion-Based Interaction for Augmented Reality Head-Mounted Displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300674

[72] Wenge Xu, Hai-Ning Liang, Yuxuan Zhao, Tianyu Zhang, Difeng Yu, and Diego Monteiro. 2019. 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 (2019), 1991–2001. https://doi.org/10.1109/TVCG.2019.2898736

[73] Hisao Yamada. 1980. *A historical study of typewriters and typing methods, from the position of planning Japanese parallels*. Vol. 2. Journal of Information Processing. 175 – 202 pages.

[74] Naoki Yanagihara, Buntarou Shizuki, and Shin Takahashi. 2019. Text Entry Method for Immersive Virtual Environments Using Curved Keyboard. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 116, 2 pages. https://doi.org/10.1145/3359996.3365026

[75] Chun Yu, Yizheng Gu, Zhican Yang, Xin Yi, Hengliang Luo, and Yuanchun Shi. 2017. Tap, Dwell or Gesture? Exploring Head-Based Text Entry Techniques for HMDs. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4479–4488. https://doi.org/10.1145/3025453.3025964

[76] Difeng Yu, Kaixuan Fan, Heng Zhang, Diego Monteiro, Wenge Xu, and Hai-Ning Liang. 2018. PizzaText: Text Entry for Virtual Reality Systems Using Dual Thumbsticks. *IEEE Transactions on Visualization and Computer Graphics* 24, 11 (2018), 2927–2935. https://doi.org/10.1109/TVCG.2018.2868581

[77] Lian Yue, Lu Zongxing, Dong Hui, Jia Chao, Liu Ziqiang, and Liu Zhoujie. 2023. How to Achieve Human-Machine Interaction by Foot Gesture Recognition: A Review. *IEEE Sensors Journal* (2023), 1–1. https://doi.org/10.1109/JSEN.2023.3285214

[78] Shumin Zhai and Per-Ola Kristensson. 2003. Shorthand Writing on Stylus Keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI '03)*. Association for Computing Machinery, New York, NY, USA, 97–104. https://doi.org/10.1145/642611.642630