Int. J. Human–Computer Studies 193 (2025) 103385

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Comparing typing methods for uppercase input in virtual reality: Modifier Key vs. alternative approaches

Min Joo Kim [a], Yu Gyeong Son [a], Yong Min Kim [b,1], Donggun Park [a,*]

[a] Department of Media Communication, Pukyong National University, Busan, Republic of Korea
[b] Division of Interdisciplinary Studies in Cultural Intelligence (HCI Science Major), Dongduk Women's University, Seoul, Republic of Korea

A R T I C L E I N F O

Keywords:
Virtual reality (VR)
Virtual environment (VE)
Typing interaction
User experience (UX)
Modifier key

A B S T R A C T

Typing tasks are basic interactions in a virtual environment (VE). The presence of uppercase letters affects the meanings of words and their readability. By typing uppercase letters on a QWERTY keyboard, the layers can be switched using a modifier key. Considering that VE controllers are typically used in a VE, this input method can result in user fatigue and errors. Thus, this study proposed new alternative interactions for the modifier key input and compared their typing performance and user experience. In an experiment, 30 participants were instructed to type 10 sentences using different typing interaction methods (shift, long press, and double-tap) on a virtual keyboard in a VE. The typing speed, error rate, and number of backspace inputs were measured to compare typing performance. Upon the completion of the typing task, the usability, workload, and sickness associated with each typing method were evaluated. The results showed that the double-tap method exhibited significantly higher typing speed, error rate, ease of use, satisfaction, and workload. This result is consistent with those of previous studies demonstrating that selection tasks were more efficient with fewer hand movements. Thus, this study implies that the double-tap method can be considered as a potential typing interaction for the VEs instead of the traditional method using the shift as a modifier key. Therefore, this study is expected to contribute to the design and development of user-friendly interactions.

1. Introduction

The global virtual reality (VR) market is expected to grow at a compound annual growth rate (CAGR) of 27.5 % by 2030 (Grand View Research 2023). New technologies such as VR, augmented reality (AR), and mixed reality (MR), along with market growth, continue to drive companies such as IKEA, Microsoft, and Zoom to expand their services to virtual environments (VE) (Greener, 2021). For instance, Zoom and Microsoft provide users with virtual offices that are accessible through VR headsets. These services enable users to access PC displays and keyboards in a virtual office environment. With the extension of services to VEs, the immersion and seamless interaction that users experience in VEs indicate the potential for enhanced user experience (UX). Consequently, VEs have gained prominence as the next-generation communication channel (Gleim et al., 2024; Jung et al., 2023).

Of the various interactions that occur in VR, typing tasks are considered among the basic interactions between the user and the system (Dudley et al., 2019). Therefore, typing systems for VR should be designed to keep users in the VR environment for longer and consequently support a wider range of input methods (Speicher et al., 2018). Text entry is typically used when sending emails, instant messages, etc. with upper/lowercase letters, numbers, and symbols. In particular, the presence (or absence) of uppercase letters can significantly affect the meaning and readability of words (Wan et al., 2024). For example, "China" implies the country, whereas "china" implies the pottery. Further, "Earth" implies the planet, whereas "earth" implies the ground or soil. Moreover, uppercase letters are used when writing the first letter of a sentence, a day of the week, months, and proper nouns, such as the name of a person or place. A combination of upper/lowercase letters is recommended as a strong password (Ma et al., 2014).

There is no standard for text entry in VEs; however, studies on user-system typing interactions in VEs are underway. In VEs based on head-mounted displays (HMD), controller pointing is considered the most common typing method (Speicher et al., 2018). The controller is recommended as a typing tool for VE owing to it being a relatively familiar interaction method for users. Controller pointing employs two

* Corresponding author at: Department of Media Communication, Pukyong National University, 45, Yongso-ro, Nam-Gu, Busan, 48513, Republic of Korea.
1 Co-corresponding author: Yong Min Kim as the co-corresponding author for equivalent contribution to the article.
E-mail address: dgpark@pknu.ac.kr (D. Park).

https://doi.org/10.1016/j.ijhcs.2024.103385
Received 4 June 2024; Received in revised form 4 October 2024; Accepted 8 October 2024
Available online 12 October 2024

controllers for typing through ray casting, which is a standard pointing technology for VE user interfaces (UI). Users control the direction of a virtual ray or laser beam with physical hand movements and select virtual objects via the intersection of the ray and objects. The application of touch interaction, used in two-dimensional (2D) displays, to VEs is plagued by the issue of requirement of users moving close to a virtual object to select it. Considering this, ray casting, which does not require the user to approach objects, is advantageous as an interaction tool in VEs. Users are accustomed to this restriction in the real world; however, appropriate methods to circumvent this in the VEs are required (Bowman et al., 2012). Therefore, the controller pointing interaction appears to be an appropriate interaction method for the VEs.

However, the controller pointing interaction also has limitations for use as a VR typing method. As small hand rotations can result in large movements at the end of a ray, the selection of small objects is challenging (Bowman et al., 2012). In contrast to smartphone interfaces, the lack of a physical display renders it difficult to perform precise movements such as dragging an object to an appointed location and selecting an object precisely. This can cause fatigue and result in errors (Lindeman et al., 1999; Bowman et al., 2012; Lee and Hui, 2018). Therefore, new typing interactions that minimize the movement of the pointer or hand to reduce user fatigue and errors must be explored.

Although various keyboard layouts have been proposed to reduce user fatigue, the most preferred and dominant text-entry paradigm is the QWERTY keyboard (Gopher and Raij, 1988). A QWERTY keyboard typically comprises multiple layers for accessing various types of characters. Users switch between these layers using modifier keys (e.g., shift and cap lock) to input the desired keys. Both physical and touchscreen keyboards switch layers use modifier keys (Wan et al., 2024). Although studies on virtual keyboards have explored the usability of input tools, those on interaction methods that consider typing contexts, such as inputting modifier keys, are lacking. In virtual keyboard studies, the predominant approach involves extending 2D UI typing interaction methods, which involves changing layers into three-dimensional environments with minimal exploration of new interaction methods. Typing in a VE utilizes controllers based on a ray-casting system, whereas traditional typing methods rely on a user's finger movements. Thus, users may encounter difficulties in completing tasks smoothly and may experience satisfaction through traditional interactions.

This study proposed novel interaction methods for virtual keyboards that used controllers to enhance the usability of modifier keys. In particular, this study compared the typing performance and UX between the traditional method of uppercase input and new interaction methods that replace the modifier key input, particularly in uppercase input scenarios. To verify this, the following hypotheses were formulated:

H1. Typing performances exhibit no difference between interaction methods for inputting uppercase letters.

H2. UX exhibit no difference between interaction methods for inputting uppercase letters.

H3. Workloads exhibit no difference between interaction methods for inputting uppercase letters.

2. Literature review

2.1. VR interaction tools

Many studies have been conducted on the input interaction between users and systems in VEs (Kim and Xiong, 2022b, 2022a; Speicher et al., 2018). Interaction tools can be classified into two categories: controllers and hand tracking. The controller interaction mimics a user's hand movements and positions. It involves various inputs, such as grabbing or manipulating virtual objects through buttons or triggers (Hameed et al., 2023). Controller pointing and tapping are examples of controller interactions. In controller pointing, users aim at keys via ray casting and press the trigger button for input (Speicher et al., 2018). Controller tapping involves flipping the controller over and using it as a pen to press keys located at the end of the controller for input (Speicher et al., 2018). A cusor is a controller-based interaction that utilizes the touchpad of the controller. This facilitates users in moving the cursor between keys by pressing a direction, and the input by pressing the trigger button (Speicher et al., 2018).

Hand-tracking interactions map real-world hand movements and positions to virtual hand gestures. This is achieved by attaching sensors to the user's hands or utilizing cameras to track hand and finger movements (Hameed et al., 2023). The hand-tracking interaction involves the free hand, pinch, haptic glove, etc.. The freehand tracks the user's actual finger movements, necessitating no separate equipment and providing natural interactions (Speicher et al., 2018). Pinching has been suggested as an alternative to haptic feedback in mid-air tasks. This offers advantages in the UX without additional equipment or costs (Kim and Xiong, 2022b). The haptic glove offers haptic feedback and tracks the user's hand movements. However, it has limitations in that the various hand shapes and sizes require additional equipment and incur high costs (Perret and Vander Poorten, 2018).

Moreover, many interaction methods have been suggested, such as eye and head tracking. Eye tracking involves the selection of keys based on the duration of gaze and utilizes muscles that are unaffected in most situations. This method offers the advantage of being used by users with various levels of motor impairment. However, there are limitations such as slow typing speed, difficulty in fine-pointing tasks, and eye fatigue owing to the need to maintain sufficient dwell time (Feng et al., 2021). Simultaneously, limitations such as slow typing speed, difficulty in fine-pointing tasks, and eye fatigue arise from the need to maintain the dwell time (Feng et al., 2021). Head tracking is a representative hands-free interaction method and involves pointing via tracking the direction of the head position or facial-specific factors such as the end of the nose (Gizatdinova et al., 2023). Although head-tracking offers a more accurate input than eye-tracking, it is slower (Gizatdinova et al., 2018).

2.2. VR and controller

Immersive and realistic VR experiences attract users. However, the sustained interest and motivation of users is dependent on the ease and intuitiveness of the interaction or interface provided by the content or task. Thus, usability is crucial (Ricca et al., 2020; Hameed et al., 2023). Similarly, while most VR interactions are mediated primarily through tools, the main goal has always been the achievement of natural and intuitive interactions (Regazzoni et al., 2018). Although natural interaction has been enabled with the advancement of the hand-tracking technology, the controller remains a crucial tool for interaction in VEs. Thus, tracking accuracy and reliability must be considered in VE interactions. However, VR HMD face limitations in detecting physical hand movements, particularly fast and complex gestures (Johnson et al., 2023). In contrast, handheld controllers are considered more accurate and reliable input tools because they do not rely on headset tracking. This controller is faster and more efficient than hand-tracking (Caggianese et al., 2019; Masurovsky et al., 2020; Rantamaa et al., 2023).

Higher similarity between the use of interaction tools and reality does not always indicate efficient tools in the VE. Thus, a realistic interaction does not always guarantee higher performance. Masurovsky et al. (2020) found that interactions without controllers received positive subjective evaluations from users; however, their performance was poor. Bowman et al. (2012) compared UX based on the realism of interactions in VEs and found that simply rendering the interactions as more faithful to the real world did not guarantee performance improvement. Ease of interaction, tracking accuracy, immersion, and realism are equally important in VEs for sustaining the use of VR.

Consequently, the controller is an accurate interaction tool for VR.

2.3. VR and typing

In contrast to a physical keyboard, this controller uses only a single point. This renders the simultaneous and rapid input multiple keys onto a virtual keyboard using a controller. Thus, pointer and hand movements must be minimized to reduce user fatigue and errors (Li et al., 2006). Previous studies on virtual keyboard layouts in VE proposed new layouts, such as grouping frequently used key pairs closely together (Herthel and Subramanian, 2020; Li et al., 2006; Wang et al., 2023a). However, the QWERTY keyboard is most familiar to users. Studies have also been conducted on the layouts of physical keyboards. Since its first proposal in 1872, the QWERTY keyboard has been argued to reduce accuracy and typing speed (Zecevic et al., 2000; Nevala-Puranen et al., 2003). Although the Dvorak keyboard was proposed in 1932 as an alternative to the QWERTY keyboard, it could not overcome the QWERTY keyboard familiarity and the cost of reeducating users about its use (Gopher & Raij, 198)). Consequently, new input interactions need to be suggested to minimize pointer and hand movements while maintaining the QWERTY keyboard layout in the VEs.

In addition, the input methods for VR interactions were also discussed. Wang et al. (2023b) proposed two input methods, tap and trace, and compared their performance and UX. The tap method exhibited a word 4.3 word per minute (WPM) of average, and the trace method was 5.9 WPM of average. Although the trace method was more creative and efficient, it increases frustration and mental workload. Thus, Li et al. (2022) proposed a sew-typing method that borrowed sewing to provide a better interaction method for multiple selection situations. It achieved fast typing speed of 26.57 WPM and received positive user evaluations for its ease of use and attractiveness.

Thus, previous studies only explored the interaction tools, layouts, and basic input methods and did not consider additional typing contexts. This study proposed new interactions in the context of VR typing and comparison of their performances and UX.

3. Method

3.1. Participants

Thirty Koreans (15 males and 15 females) aged between 20 and 30 years (M = 23.2, SD=16) and 156–183 cm (M = 167.8, SD=8.5) participated in this experiment. All the participants had normal or corrected-to-normal vision. Twenty-two participants had previous experience with VR HMDs and 18 of them had experience of approximately once a year. All participants provided informed consent for the experimental protocol that was approved by the Pukyong National University Institutional Review Board (IRB NO.: 1,041,386–202,310-HR-116–02).

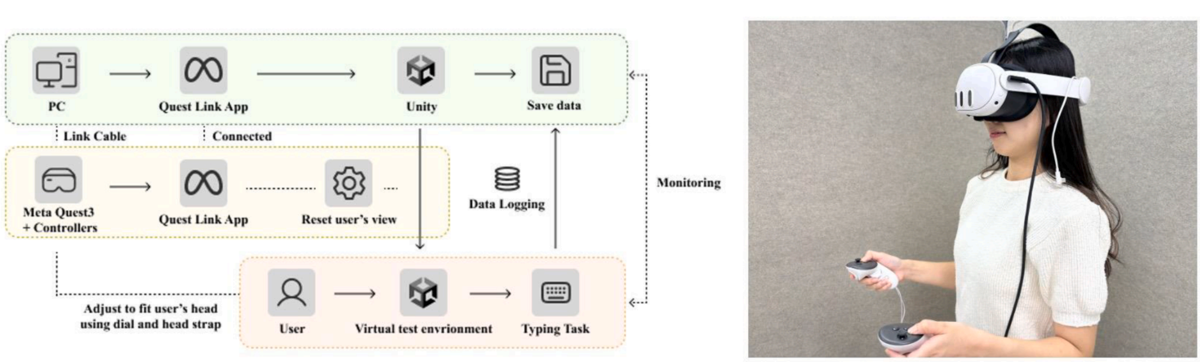
3.2. Experimental settings

The participants were equipped with Meta Quest3 (resolution: 2064 $\times$ 2208 pixels per eye) and used Touch Plus controllers of Meta Quest3. The schematic of the equipment setup and procedure is presented in Fig. 1(a). A PC was used to run the Quest Link App, which connected the Meta Quest 3 VR headset and controllers to the system. The VR headset was connected to the PC via a compatible USB 3.2 Type C Link cable. The PC was equipped with an Intel Core i9–13,900 processor running at 3.2 GHz, 64 GB of RAM, and an NVIDIA GeForce RTX 4090 GPU running Windows 11. The virtual test environment, where the typing tasks were conducted, was developed using Unity 2021.3.16f1, and all task data were saved directly to the PC for subsequent analysis. The Meta Quest 3 headset, along with its controllers, was used by the participants to interact with the VE. The Quest Link App not only enabled the connection between the VR headset and the VE but also facilitated real-time adjustments, allowing the system to reset the user's view within the virtual space. The system recorded various raw data related to input, such as time, typed text, and the number of backspaces pressed. The VR headset worn by a participant and the physical environment of the experiment are shown in Fig. 1(b).

In the virtual test environment, the virtual keyboard was located 0.35 m below the height of the VR headset (Yang et al., 2019). Here the VE, the area within a 0.5 m radius from the user, is referred to as a no-no-zone. Placing content in these areas may cause user resistance (Alger, 2015). Therefore, the initial distance between the user and the keyboard was set to 0.6 m (Fig. 2). During the practice tasks, participants were allowed to step back freely to determine the optimal horizontal distance from the keyboard. The participants performed the typing tasks approximately 0.22 m apart from the initial position.

3.3. Experimental design

Three typing methods (*shift, long press*, and *double tap*) were developed, and their typing performance and usability were evaluated (Fig. 3). The *Shift* method mirrored the conventional uppercase input method for smartphones, wherein users activated the shift key prior to selecting the desired letter. In the *long press (LP)* method, users can type uppercase letters by holding a key for >0.5 s. The *double tap (DT)* facilitates a method for inputting an uppercase letter by pressing a key twice. The interval between the first and second inputs was 0.5 s. The time parameters for *LP* and *DT* were determined based on an internal pilot test based on trial and error (Thorndike, 1933). The controllers were used to type the texts using all the typing methods. The participants aimed for a key with a virtual ray and input text by pushing the trigger button. Consequently, the system exhibited a dark background



(a) A schematics of the equipment setup and procedure **(b) Physical environment of the experiment**

Fig. 1. Physical test environmental settings. (a) A schematic of the equipment setup and procedure, (b) Physical environment of the experiment.

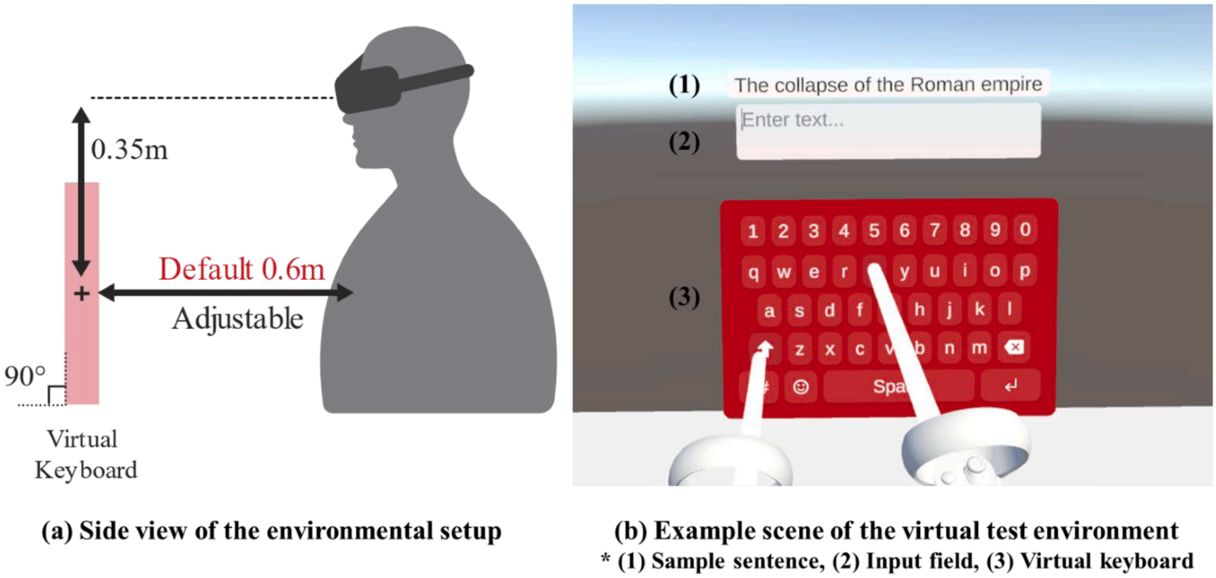


Fig. 2. Virtual test environmental settings. (a) Side view of the environmental setup, (b) Example scene of the virtual test environment.

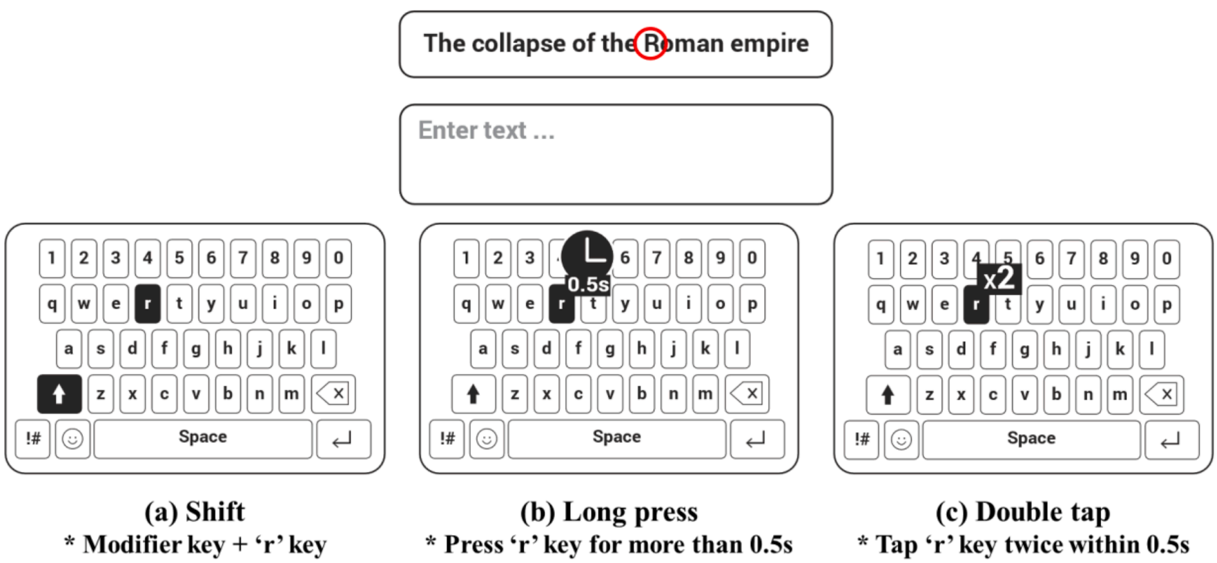


Fig. 3. Typing methods (a) Shift, (b) Long press, (c) Double tap.

color (B72A30), when the key was aimed at and a darker background color (A1252A) when the key was pressed, providing the user with visual feedback on their input.

3.4. Measures

3.4.1. Performance measures

Typing performance metrics, WPM, and the uncorrected error rate (UER), the number of backspaces were measured. WPM was computed by following Eq. (1). Further, S represents the time elapsed between the first and last key presses, and T represents the length of the transcribed text (Dudley et al., 2019; Cui and Mousas, 2023; Wang et al., 2023b).

$$\mathrm{WPM} = \frac{|T-1|}{\mathrm{s}} \mathrm{x} \frac{60}{5} \tag{1}$$

The UER was computed using the following Eq. (2), where C represents the number of correct characters in the transcribed text. Further, INF represents the minimum string distance (Soukoreff and MacKenzie, 2001), and IF is the number of deleted characters in the entire sequence (Soukoreff and MacKenzie, 2001; Dudley et al., 2019; Kim and Xiong, 2022b; Zhang and Wobbrock, 2019).

$$\mathrm{UER} = \frac{\mathrm{INF}}{\mathrm{C} + \mathrm{INF} + \mathrm{IF}} \tag{2}$$

3.4.2. Subjective experience measures

Subjective experiences were assessed using the USE, NASA-TLX, and virtual reality sickness questionnaire (VRSQ). The USE questionnaire assessed usefulness, ease of use, ease of learning, and satisfaction (Lund, 2001). The NASA-TLX is a widely used questionnaire for subjective evaluation. It measures mental, physical, temporal demands, as well as performance, effort, and frustration levels (Hart, 2006). The VRSQ evaluates VR sickness focusing on oculomotor function and disorientation (Kim et al., 2018). The items are listed in Tables 1–3, and the measured VRSQ scores were calculated using the methodology outlined in Table 4.

Table 1
USE questionnaire aspects and items.

Aspect	Item
Usefulness	It helps me be more effective.
Usefulness	It helps me be more productive.
Usefulness	It is useful.
Usefulness	It gives me more control over the activities in my life.
Usefulness	It makes the things I want to accomplish easier to get done.
Ease of use	It is easy to use.
Ease of use	It is simple to use.
Ease of use	It is user friendly.
Ease of use	It requires the fewest steps possible to accomplish what I want to do with it.
Ease of learning	I learned to use it quickly.
Ease of learning	I easily remember how to use it.
Ease of learning	It is easy to learn to use it.
Satisfaction	I would recommend it to a friend.
Satisfaction	It is fun to use.
Satisfaction	It works the way I want it to work.

Note. Each item is rated based on Likert 7-point scale (1 = strongly disagree, 7 = strongly agree).

Table 2
NASA-TLX questionnaire aspects and items.

Aspect	Item
Mental demand	How much mental and perceptual activity was required?
Physical demand	How much physical activity was required?
Temporal demand	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred?
Performance	How successful do you think you were in accomplishing the goals of the task?
Effort	How hard did you have to work to accomplish your level of performance?
Frustration	How insecure, discouraged, irritated stressed and annoyed did you feel during the task?

Note. Each dimension is rated on a scale from 0 to 100 with increments of 5.

Table 3
VRSQ symptoms and simulator sickness questionnaire (SSQ) components.

No.	Symptom	SSQ components
1	General discomfort	Oculomotor
2	Fatigue	Oculomotor
3	Eyestrain	Oculomotor
4	Difficulty focusing	Oculomotor
5	Headache	Disorientation
6	Fullness of head	Disorientation
7	Blurred Vision	Disorientation
8	Dizzy (eyes closed)	Disorientation
9	Vertigo	Disorientation

Note. Each item is rated with the scale of 0 to 3 (0 = none, 1 = slight, 2 = moderate, 3 = severe).

Table 4
Computation score of VRSQ.

SSQ components	Computation
Oculomotor	{([1]+[2]+[3]+[4])/12}*100
Disorientation	{([5]+[6]+[7]+[8]+[9])/15}*100
Total	(Oculomotor score + Disorientation score)/2

3.5. Experimental procedure

Fig. 4 illustrates the experimental procedure. Participants completed practice tasks before the main experiment. The participants adapted each typing method *(Shift, LP, DT)* in a practice session lasting approximately 7 min. During the practice session, participants typed three sentences provided using each method and were instructed to step back appropriately for typing. The participants performed the tasks in their respective positions they had set. They typed text while equipping the HMD with the randomly presented methods. The participants were instructed to type 10 sentences as quickly and accurately as possible using each typing method. Thus, the 10 sentences were composed of a mixture of English uppercase and lowercase letters. Thirty sentences were provided to the participants, all from the phase sets of (MacKenzie and Soukoreff, 2003) and Enron Mobile (Vertanen and Kristensson, 2011). They comprised more than two uppercase letters and fewer than 43 letters. The sentences were 280–302 characters in length, including spaces, and contained 29–30 capitalized letters across the ten sentences. Under the *DT* condition, six words such as "see," "will," and "agree," where the same lowercase letter appeared consecutively, were included across the ten presented sentences. After completing the test under each condition, the participants removed their HMD and filled in the questionnaires, followed by 2 min rest. Once the tests and subjective evaluations for all typing methods were completed, the participants were asked to choose their preferred method. The entire experimental procedure lasted approximately 45 min. Except for three sentences wherein the enter key was pressed before the participants finished typing, 897 sentences were analyzed.

3.6. Data analysis

This study investigated data on the timestamps (in ms) of the initial input and enter key inputs, the total number of backspace key inputs in a sequence, sample text, and the text finally entered by the participant. The text input time was the duration between the first and last keypresses. The minimum string distance was computed through comparisons of the presented sample and final transcribed text data. Python code was used to process all the data.

One-way repeated-measures analysis of variance (RM-ANOVA) and Bonferroni tests for post hoc grouping were performed to check the statistical significance of each response. Statistical analyses were conducted using IBM SPSS Statistics version 27 at a significance level of $p <$ .05. The degrees of freedom were corrected using the Greenhouse-Geisser correction if the epsilon from Mauchly's test of sphericity was below 0.75, and the Huynh-Feldt correction if the epsilon was above 0.75.

4. Result

4.1. Typing performance

Fig. 5 shows the text entry performance metrics: WPM, UER, and the number of backspaces. There was a statistically significant difference between the *LP* and *DT* in WPM ($F = 7.275$, p=.003, η^2=0.201). The *DT* ($M = 8.831$, SD=1.52) was the fastest keyboard and the *LP* ($M = 7.759$, SD=2.36) was the slowest. Although no statistically significant difference was observed in UER, the *Shift* ($M = 2.076$, SD=4.01) showed the highest error rate, and the *DT* ($M = 1.368$, SD=2.52) exhibited the lowest error rate. The number of backspaces was statistically significant in all keyboard types ($F = 33.685$, p=.000, η^2=0.537). The backspace inputs were in the *LP* ($M = 4.933$, SD=2.37), *DT* ($M = 3.144$, SD=1.32), and the *Shift* ($M = 2.010$, SD=1.11). In conclusion, the *DT* demonstrated faster speed and fewer errors compared to other methods. This suggests that the *DT* can support more efficient typing tasks in VEs than other methods.

4.2. Subjective experience evaluation

The USE questionnaire was used to evaluate the UX. The differences observed owing to the uppercase input method were statistically

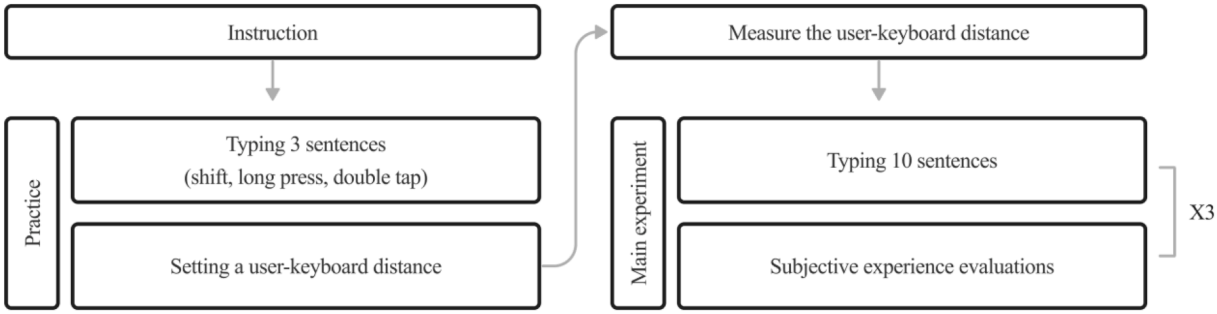


Fig. 4. Experimental procedure.

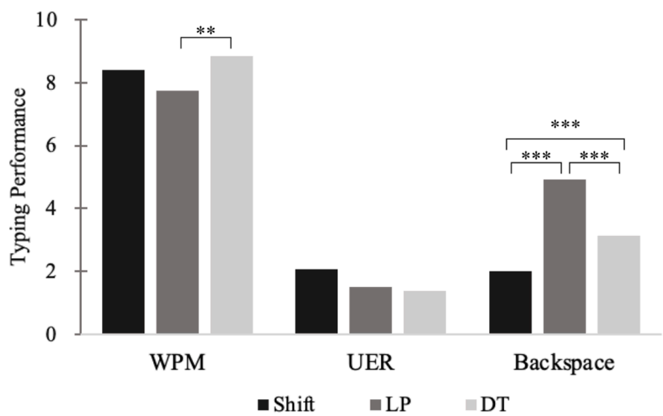


Fig. 5. Comparison results of typing performance. WPM (Counts), UER (Percentage), Backspace (Counts). *** indicates significant difference with $p < .001$, ** indicates significant difference with $p < .01$, * indicates significant difference with $p < .05$.

significant in usefulness (p=.000, F = 10.392, η^2=0.264), ease of use (F = 4.058, p=.022, η^2=0.123), ease of learning (F = 5.868, p=.005, η^2=0.168), and satisfaction (F = 7.325, p=.001, η^2=0.202) (Fig. 6). The post-hoc grouping exhibited that the *Shift* (M = 5.553, SD=1.12) and the *DT* (M = 5.273, SD=1.52) were more useful than the *LP* (M = 4.187, SD=1.73). Here, the *DT* (M = 5.700, SD=1.21) was the easiest to use, and the *Shift* (M = 6.711, SD=0.49) was the easiest to learn. Further, the *Shift* (M = 5.392, SD=1.27) and the *DT* (M = 5.417, SD=1.54) satisfied users more than the *LP* (M = 4.208, SD=1.63). Overall, the *LP* was evaluated as having a diminished UX compared to the other two methods. This result appears to stem from differences in factors such as familiarity, feedback mechanisms, and input frequency inherent to each interaction method.

The VRSQ was used to evaluate sickness after typing tasks. Although there was no statistically significant difference between the typing methods, the sickness was observed in the *Shift* (M = 9.99, SD=12.86), *DT* (M = 8.94, SD=10.64), and *LP* (M = 8.14, SD=9.47) order. The VRSQ, developed based on the SSQ to assess sickness in VR environments using a 100-point scale, indicates that a lower score corresponds to either negligible or minimal sickness symptoms (Kim et al., 2018). According to the analysis, there was no statistically significant difference between the three methods, and all mean scores were below 10, suggesting that the three typing methods have minimal impact on VR sickness.

The *DT* (M = 15.36, SD=13.79) was significantly lower than the *LP* (M = 22.61, SD=16.56) in case of the NASA-TLX score (F = 7.942, p=.003, η^2=0.215). Further, the *DT* was significantly lower than the *LP* in terms of mental demand (F = 8.068, p=.004, η^2=0.218), effort (F = 3.844, p=.027, η^2=0.117), and frustration (F = 8.644, p=.003, η^2=0.230) and was significantly higher than the *LP* in terms of performance (F = 9.779, p=.000, η^2=0.252) (Fig. 7). In conclusion, the *LP* required a generally higher workload compared to the other methods. This result suggests that the *Shift* and *DT* have greater potential as more efficient interactions for typing in VEs.

4.3. Error rate by VR experience

This study investigated the typing error rate according to users' previous experiences with VR. An independent t-test was conducted by dividing the participants into inexperienced and experienced groups. The inexperienced group comprised 14 participants with one or less VR

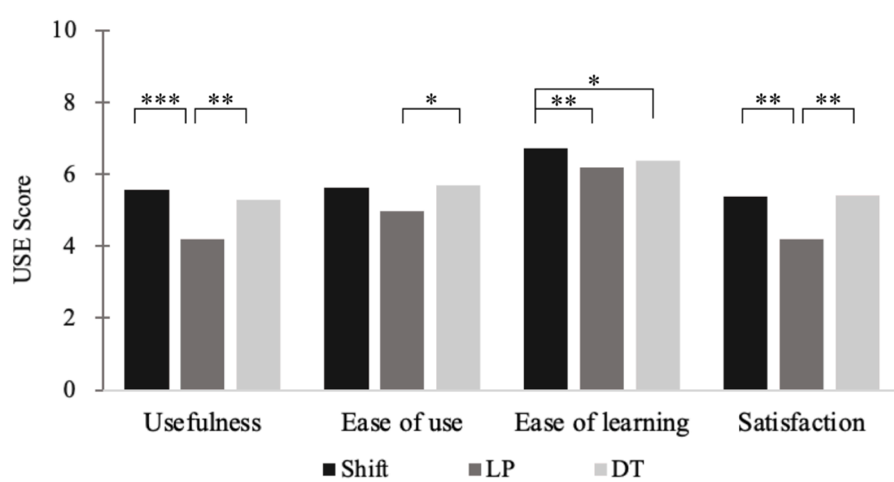


Fig. 6. Comparison results of USE. *** indicates significant difference with $p < .001$, ** indicates significant difference with $p < .01$, * indicates significant difference with $p < .05$.

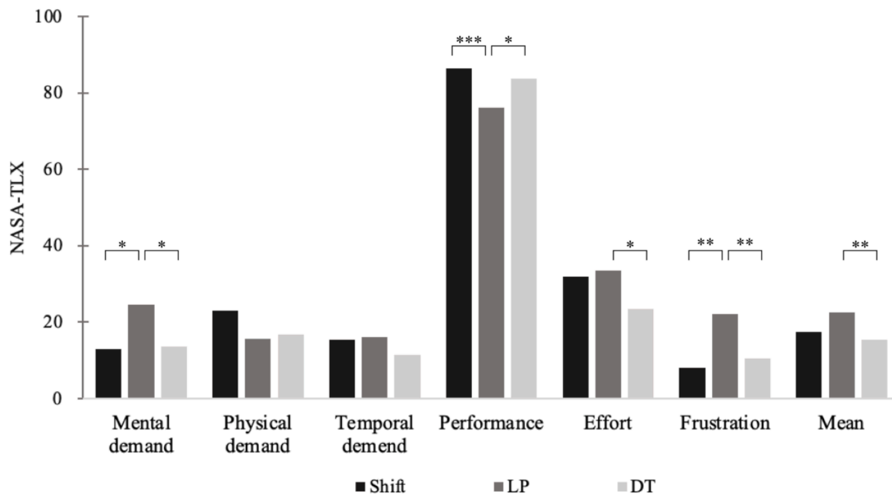


Fig. 7. Comparison results of NASA-TLX score. *** indicates significant difference with $p < .001$, ** indicates significant difference with $p < .01$, * indicates significant difference with $p < .05$
*Note: The Performance score was calculated by subtracting the value from 100 to obtain the mean value.

experiences. Whereas, the experienced group comprised 16 participants with more than one VR experience. Although there was no significant difference, the experienced group exhibited a lower error rate than the inexperienced group for all typing methods (Table 5).

4.4. Preference

After completing all the typing tasks, the participants were asked to respond with their most preferred method and the reason for their choice (Table 6). Eighteen participants selected the *DT* as their preferred typing method for uppercase input, followed by the *Shift* (9 participants) and *LP* (3 participants) methods($n = 30$). In the follow-up interview, participants who chose *DT* addressed the reasons for their choice as follows: low physical demand and speed (10 cases), low error rate and ease of control (5 cases), and fun and pleasure when inputting uppercase letters (3 cases). The participants who selected *Shift* described it as "the most familiar method" (6 cases) and "precise for typing with my intention" (3 cases). The participants who chose *LP* described it as "the simplest method due to its low click frequency" (3 cases).

5. Discussion

Certain differences were observed in the typing performance according to typing interaction methods. The *DT* method was significantly faster than the *LP* method. The number of backspaces exhibited significant differences between the three methods, thereby rejecting the first hypothesis. The typing interaction methods also exhibited differences in UX and workload. The *DT* method was significantly better in terms of ease of use and satisfaction than the *LP* method. The *DT* method also resulted in a lower workload in terms of mental demand, effort, and frustration. Thus, the second and third hypotheses were rejected. In particular, the *DT* method exhibited a lower typing error rate and perceived workload than the traditional *shift* method according to the mean of the measured scales. It also exhibited higher typing speed, subjective evaluations, and preferences. The *DT* method offered advantages in terms of typing speed, error rate, ease of use, satisfaction, and workload compared to the *shift* method. The results of this study showed that the *DT* method exerted positive effects on the typing performance, UX, and preference for VR typing using controllers.

Table 5
Independent *T*-test results for Error rate (UER) by VR experience.

Typing methods	Experienced		Inexperienced		*t*	*df*	*p*
	Mean	SD	Mean	SD			
Shift	1.476	3.043	2.765	4.932	0.874	28	0.389
LP	0.763	1.362	2.386	3.519	1.708	28	0.099
DT	1.036	1.275	1.749	3.464	0.767	28	0.450

Table 6
Results of the preference and the user interview.

Typing methods	*N*	Reasons for preference	*N*	Examples of responses
Shift	9	Familiarity	6	"It was familiar and comfortable since it was the most similar method to what I usually use." (P2)
		Precision	3	*"There were fewer instances where uppercase letters were entered unintentionally. When I pressed Shift, uppercase letters appeared on the keyboard, making it clear that it was in an uppercase input state."* (P23)
LP	3	Simplicity	3	*"It was the simplest way without the need to press additional keys or move my hands more."* (P22)
DT	18	Low physical demand & speed	10	*"The distance my hands needed to move in mid-air was short, which reduced physical demand, and I could type the fastest since there were no cumbersome movements."* (P15)
		Ease of control & error rate	5	*"It felt like a natural typing method in a virtual environment. I felt in control of my input, and there were the fewest typing errors."* (P27)
		Fun	3	*"The rhythm of double-tapping made it more enjoyable than the familiar typing method (Shift)."* (P28)

The *Shift* method received the highest evaluation in terms of usefulness and ease of learning. This method switches layers using a modifier key, yielding uppercase letters on the UIs of the keys. This rendered it easier to prevent user mistakes compared to other methods. However, the *DT* and *LP* methods did not switch layers to input the upper letters. The users described the *DT* method as "I should always be aware that continuous inputs of lowercase letters result in the uppercase letters. If not, an unintended letter was input." Although users can utilize the *Shift* method without learning, they must learn and use the *DT* method. This demonstrated the characteristics of the new typing interactions (Wang et al., 2023b). As indicated by the user interviews, a good evaluation of the *shift* method originated from its characteristic of being the most used method. This result was consistent with prior studies showing that interaction experiences and familiarity moved to a new environment, thereby lowering mental demands (Jacob et al., 2007; NORMAN, 2011). This is also consistent with the results of a study wherein experiences helped users clearly understand functionalities, enabling them to perceive purpose and usage without additional learning processes (Hollinworth and Hwang, 2011).

However, this study focused on the determining the results of the *DT* method that achieved the best evaluation in terms of ease of use, satisfaction, preference, perceived workload, and typing performance. This result implies that the advantages of the *DT* keyboard performance and UX overcome the cost of teaching users new interactions (Boletsis and Kongsvik, 2019; Wang et al., 2023b). The *DT* method demonstrated the lowest workload in terms of effort and overall NASA-TLX scores. As stated in 3.5. Experimental procedure, there were cases that the same lowercase letter appeared consecutively under the *DT* condition. This may lead to unintended inputs contrary to user intent. Despite this issue, the *DT* method demonstrated good performance. The *DT* method was faster (8.83 WPM) than the other methods, and its difference from the *LP* method was statistically significant. The *DT* method exhibited the lowest error rate. These results were related to the characteristics of the VE interactions. Bowman (2002) referred to two tasks in VR interactions: selection and manipulation. The *DT* method inputs uppercase letters without aiming at the two keys, thereby skipping the modifier key selection task. Thus, the *DT* method offers the advantage of enabling an uppercase input in fewer steps. Interactions in VR rely on interaction tools that track user movements in three dimensions. This results in inaccurate interactions owing to the noise caused by natural hand tremors (Weise et al., 2020). Considering this, it can be inferred that the *DT* method, which requires minimal movement for accurate key targeting in the VEs, exhibited high performance. Although the number of inputs in the *DT* method was the same as that in the *Shift* method, the movements required for inputting were smaller in the *DT* method. This result demonstrated that the *DT* method exhibited superior performance and UX, aligning with Fitts' law (Fitts, 1954), which posits that longer pointer movements require more time to complete. This can be predicted from a previous study (Poupyrev et al., 1997) in which the efficiency of the controller pointing interaction increases when hand movements are minimized, particularly in the continuous selection of small objects in VEs. Although this study reported great performance and UX with the *DT*, there were two cases reported that uppercase letters were input unintentionally during consecutive inputs in the experiment throughout all trials. Therefore, an option for users to customize the interval time should be considered, and future study needs to closely explore the optimal interval time.

The *LP* method also enables capitalization with minimal hand movement and is one of the most intuitive interactions, with the fewest input steps and physical demands. However, despite these advantages, the *LP* method received relatively low ratings for typing speed and UX. This was consistent with previous studies (Bowman et al., 2012; Masurovsky et al., 2020) suggesting that the intuitiveness of interactions does not always correlate linearly with performance or utility. Therefore, this indicates that higher intuitiveness in interactions does not always result in higher performance or utility. Moreover, the *LP* method is prone to frequent mistakes by users. A long press is a gesture wherein the duration of the touch exceeds the time threshold after the initial contact (Quinn et al., 2021). When using interactions involving dwell time, such as a long press, users must point to a specific area for a period. According to Argelaguet and Andujar (2013), in case of an excessively short duration, the system may mistakenly interpret users' actions as intentional interactions, leading to the execution of functions known as the Midas touch effect, even when unintended by the user. However, for longer durations, the need to point to a narrow area increases hand tremor, resulting in decreased accuracy (Steed, 2006). These results are consistent with those of previous studies (Argelaguet and Andujar, 2013; Steed, 2006; Kulik et al., 2012; Shin et al., 2015) were observed. The results show that the *LP* method has the highest number of backspaces within a sequence. Therefore, the option for users to customize the pressing duration for the practical implementation of the *LP* method must be considered as an alternative to the modifier key in VEs.

Controllers are the most common tools for typing in VEs, offering users a familiar way to interact. This study explored alternative interaction methods for typing in VEs, particularly focusing on enhancing the UX and minimizing physical demands in the context requiring the uppercase lowercase letters. In the comparison of the traditional *shift* modifier key with the *LP* and *DT* methods the *DT* showed promising results in certain performance metrics like typing speed and error rate, and UX measures. However, the results of this study do not conclusively demonstrate that the *DT* is always superior. Rather, the study aimed to explore new alternatives to the *shift* modifier key, presenting various interaction methods that could be beneficial depending on the context.

The practical implications of these findings are particularly relevant in scenarios where both uppercase and lowercase letters are frequently used, such as composing emails or participating in chat sessions within VEs. For example, the *DT* and *LP* method can reduce the physical effort required to switch between cases, offering a more streamlined experience compared to the traditional shift key. Additionally, this method could be advantageous when entering special characters that are typically located above number keys, as it eliminates the need for layer switching via the *shift* key. Furthermore, the alternative methods explored in this study show potential for application in non-English languages, such as Korean, where the methods could simplify the input of double consonants without requiring complex key combinations.

Ultimately, the findings of this study underscore the importance of exploring multiple interaction methods rather than relying solely on established practices like the shift modifier key. The potential for more flexible and efficient typing interactions in VEs opens up new possibilities for enhancing UX, especially in tasks requiring frequent switching between cases or symbols. These insights pave the way for future research that can further validate these methods across different contexts and user needs.

6. Conclusion

This study proposed new interactions as alternatives to modifier key inputs for layer transition in the typing context to minimize pointers and hand movements in the VE. The *Shift, LP*, and *DT* methods were developed and evaluated for their performance and UX. The results showed that the *Shift* was advantageous in terms of usefulness and ease of learning. The *DT* method outperformed the other methods in terms of typing performance, ease of use, satisfaction, and preferences. The results implied that the *DT* method has potential as a new typing interaction in the VE owing to its advantages in performance and UX. Although the *LP* method demonstrated relatively lower performance in this study, it is important to note that various user postures were not considered, and the focus was limited to typing with the controller. Future research should investigate the applicability of the *LP* method in different postures, such as one-handed typing, and with alternative input methods, including the use of hand gestures. This study also did not consider the effects of color, haptics, or auditory feedback. These issues should be comprehensively investigated in future studies.

While studies on virtual keyboards typically focus on typing tool performance and UX, this study differs in considering the typing context that users can experience. In particular, the *LP* and *DT* methods have been suggested as alternatives to modify key input interactions for layer transitions. Their performance and UX were explored comprehensively. This study found that the *DT* method has clear UX and performance benefits over the *Shift* method and that users are accustomed to it. This study is expected to contribute to user-friendly interaction design in VEs.

CRediT authorship contribution statement

Min Joo Kim: Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Conceptualization. **Yu Gyeong Son:** Writing – review & editing, Investigation, Data curation. **Yong Min Kim:** Writing – review & editing, Supervision, Methodology. **Donggun Park:** Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgments

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICAN(ICT Challenge and Advanced Network of HRD) support program supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation) (IITP-2023-RS-2023–00259806). This research was also supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT) (RS-2023–00242528).

Data availability

Data will be made available on request.

References

Alger, M., 2015. Visual Design Methods For Virtual Reality. Ravensbourne. http://aperturesciencellc.com/vr/VisualDesignMethodsforVR_MikeAlger.pdf.

Argelaguet, F., Andujar, C., 2013. A survey of 3D object selection techniques for virtual environments. Comput Graph 37 (3), 121–136.

Boletsis, C., Kongsvik, S., 2019. Text input in virtual reality: a preliminary evaluation of the drum-like vr keyboard. Technologies 7 (2), 31.

Bowman, D.A., 2002. Principles for the design of performance-oriented interaction techniques. Handbook of Virtual Environments. CRC Press, pp. 317–340.

Bowman, D.A., McMahan, R.P., Ragan, E.D., 2012. Questioning naturalism in 3D user interfaces. Commun. ACM 55 (9), 78–88.

Caggianese, G., Gallo, L., Neroni, P., 2019. The vive controllers vs. leap motion for interactions in virtual environments: a comparative evaluation. In: Intelligent Interactive Multimedia Systems and Services: Proceedings of 2018 Conference 11. Springer International Publishing, pp. 24–33.

Cui, D., Mousas, C., 2023. Exploring the effects of virtual hand appearance on midair typing efficiency. Comput Animat Virtual Worlds 34 (3–4), e2189.

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

Feng, W., Zou, J., Kurauchi, A., Morimoto, C.H., Betke, M., 2021. Hgaze typing: head-gesture assisted gaze typing. In: ACM Symposium on eye tracking research and applications, pp. 1–11.

Fitts, P.M., 1954. The information capacity of the human motor system in controlling the amplitude of movement. J. Exp. Psychol. 47 (6), 381.

research, Grand view, 2023. Virtual Reality Market Size, Share & Trends Analysis Report By Technology (Semi & Fully Immersive, Non-Immersive), By Device (HMD, GTD), By Component (Hardware, Software). April. By Application, By Region, And Segment Forecasts, 2023 –2030. Grand view research. (2024, March 26). https://www.grandviewresearch.com/industry-analysis/virtual-reality-vr-market.

Gizatdinova, Y., Špakov, O., Tuisku, O., Turk, M., Surakka, V., 2018. Gaze and head pointing for hands-free text entry: applicability to ultra-small virtual keyboards. In: Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications, pp. 1–9.

Gizatdinova, Y., Špakov, O., Tuisku, O., Turk, M., Surakka, V., 2023. Vision-Based Interfaces for Character-Based Text Entry: Comparison of Errors and Error Correction Properties of Eye Typing and Head Typing. Advances in Human-Computer Interaction, p. 2023.

Gleim, M., McCullough, H., Ferrell, O., Gabler, C., 2024. Metaverse: shifting the reality of services. J. Services Market. 38 (1), 13–27.

Gopher, D., Raij, D., 1988. Typing with a two-hand chord keyboard: will the qwerty become obsolete? IEEE Trans. Syst. Man Cybern. 18 (4), 601–609.

Hameed, A., Möller, S., Perkis, A., 2023. How good are virtual hands? Influences of input modality on motor tasks in virtual reality. J. Environ. Psychol. 92, 102137.

Hart, S.G., 2006. NASA-task load index (NASA-TLX); 20 years later. In: Proceedings of the human factors and ergonomics society annual meeting. Sage CA: Los Angeles, CA, 50. Sage publications, pp. 904–908.

Herthel, A.B., Subramanian, A., 2020. Optimizing single-finger keyboard layouts on smartphones. Comput Oper Res 120, 104947.

Hollinworth, N., Hwang, F., 2011. Investigating familiar interactions to help older adults learn computer applications more easily. In: Proceedings of HCI 2011 The 25th BCS Conference on Human Computer Interaction. BCS Learning & Development.

Jacob, R.J., Girouard, A., Hirshfield, L.M., Horn, M.S., Shaer, O., Solovey, E.T., Zigelbaum, J., 2007. Reality-based interaction: unifying the new generation of interaction styles. In: CHI'07 extended abstracts on Human factors in computing systems, pp. 2465–2470.

Johnson, C.I., Fraulini, N.W., Peterson, E.K., Entinger, J., Whitmer, D.E., 2023. Exploring Hand Tracking and Controller-Based Interactions in a VR Object Manipulation Task. In: International Conference on Human-Computer Interaction. Cham: Springer Nature Switzerland, pp. 64–81.

Jung, T., Cho, J., Han, D.-I.D., Ahn, S.J.G., Gupta, M., Gopal, D., Heo, C.Y., Loureiro, S.M. C., Sigala, M., Trunfio, M., 2023. Metaverse for service industries: future applications, opportunities, challenges and research directions. Comput Human Behav, 108039.

Kim, H.K., Park, J., Choi, Y., Choe, M., 2018. Virtual reality sickness questionnaire (VRSQ): motion sickness measurement index in a virtual reality environment. Appl. Ergon. 69, 66–73.

Kim, W., Xiong, S., 2022a. Pseudo-haptic button for improving user experience of mid-air interaction in VR. Int. J. Hum. Comput. Stud. 168, 102907.

Kim, W., Xiong, S., 2022b. Pseudo-haptics and self-haptics for freehand mid-air text entry in VR. Appl. Ergon. 104, 103819.

Kulik, A., Dittrich, J., Froehlich, B., 2012. The hold-and-move gesture for multi-touch interfaces. In: Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services, pp. 49–58.

Lee, L.-H., Hui, P., 2018. Interaction methods for smart glasses: a survey. IEEE access 6, 28712–28732.

Li, Y., Chen, L., Goonetilleke, R.S., 2006. A heuristic-based approach to optimize keyboard design for single-finger keying applications. Int. J. Ind. Ergon. 36 (8), 695–704.

Li, Y., Sarcar, S., Kim, K., Tu, H., Ren, X., 2022. Designing successive target selection in virtual reality via penetrating the intangible interface with handheld controllers. Int. J. Hum. Comput. Stud. 165, 102835.

Lindeman, R.W., Sibert, J.L., Hahn, J.K., 1999. Hand-held windows: towards effective 2D interaction in immersive virtual environments. In: Proceedings IEEE Virtual Reality (Cat. No. 99CB36316). IEEE, pp. 205–212.

Lund, A.M., 2001. Measuring usability with the use questionnaire12. Usability Interface 8 (2), 3–6.

Ma, J., Yang, W., Luo, M., Li, N., 2014. A study of probabilistic password models. In: 2014 IEEE Symposium on Security and Privacy. IEEE, pp. 689–704.

MacKenzie, I.S., Soukoreff, R.W., 2003. Phrase sets for evaluating text entry techniques. In: CHI'03 extended abstracts on Human factors in computing systems, pp. 754–755.

Masurovsky, A., Chojecki, P., Runde, D., Lafci, M., Przewozny, D., Gaebler, M., 2020. Controller-free hand tracking for grab-and-place tasks in immersive virtual reality: design elements and their empirical study. Multimodal Technol. Interaction 4 (4), 91.

Nevala-Puranen, N., Pakarinen, K., Louhevaara, V., 2003. Ergonomic intervention on neck, shoulder and arm symptoms of newspaper employees in work with visual display units. Int. J. Ind. Ergon. 31 (1), 1–10.

NORMAN, D., 2011. Living With Complexity., 2011. The Mit Press.

Perret, J., Vander Poorten, E., 2018. Touching virtual reality: a review of haptic gloves. In: ACTUATOR 2018; 16th International Conference on New Actuators. VDE, pp. 1–5.

Poupyrev, I., Weghorst, S., Billinghurst, M., Ichikawa, T., 1997. A framework and testbed for studying manipulation techniques for immersive VR. In: Proceedings of the ACM symposium on Virtual reality software and technology, pp. 21–28.

Quinn, P., Feng, W., Zhai, S., 2021. Deep Touch: Sensing Press Gestures from Touch Image Sequences. Artificial Intelligence for Human Computer Interaction: A Modern Approach, pp. 169–192.

Rantamaa, H.-R., Kangas, J., Kumar, S.K., Mehtonen, H., Järnstedt, J., Raisamo, R., 2023. Comparison of a vr stylus with a controller, hand tracking, and a mouse for object manipulation and medical marking tasks in virtual reality. Appl. Sci. 13 (4), 2251.

Regazzoni, D., Rizzi, C., Vitali, A., 2018. Virtual reality applications: guidelines to design natural user interface. In: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 51739. American Society of Mechanical Engineers. V01BT02A029.

Greener, Rory, 2022. Zoom, Microsoft Team Up For VR Video Conferencing. XR Today. https://www.xrtoday.com/mixed-reality/zoom-microsoft-team-up-for-vr-video-conferencing/.

Ricca, A., Chellali, A., Otmane, S., 2020. Influence of hand visualization on tool-based motor skills training in an immersive VR simulator. In: 2020 IEEE international symposium on mixed and augmented reality (ISMAR). IEEE, pp. 260–268.

Shin, H., Lim, J.M., Oh, C., Kim, M., Jeong, H.T., Son, J., 2015. Performance comparison of tap gestures on small-screen touch devices. In: 2015 IEEE International Conference on Consumer Electronics (ICCE). IEEE, pp. 120–121.

Soukoreff, R.W., MacKenzie, I.S, 2001. Measuring errors in text entry tasks: an application of the Levenshtein string distance statistic. In: CHI'01 extended abstracts on Human factors in computing systems, pp. 319–320.

Speicher, M., Feit, A.M., Ziegler, P., Krüger, A., 2018. Selection-based text entry in virtual reality. In: Proceedings of the 2018 CHI conference on human factors in computing systems, pp. 1–13.

Steed, A., 2006. Towards a general model for selection in virtual environments. In: 3D User Interfaces (3DUI'06). IEEE, pp. 103–110.

Thorndike, E.L., 1933. A proof of the law of effect. Science 77 (1989), 173–175.

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

Wan, T., Wei, Y., Shi, R., Shen, J., Kristensson, P.O., Atkinson, K., Liang, H.-N., 2024. Design and evaluation of controller-based raycasting methods for efficient alphanumeric and special character entry in virtual reality. IEEE Trans Vis Comput Graph.

Wang, K., Yan, Y., Zhang, H., Liu, X., Wang, L., 2023a. Eye-shaped keyboard for dual-hand text entry in virtual reality. Virtual Reality Intellig. Hardware 5 (5), 451–469.

Wang, Y., Wang, Y., Li, X., Zhao, C., Ma, N., Guo, Z., 2023b. A comparative study of the typing performance of two mid-air text input methods in virtual environments. Sensors 23 (15), 6988.

Weise, M., Zender, R., Lucke, U., 2020. How can I grab that? Solving issues of interaction in VR by choosing suitable selection and manipulation techniques. i-com 19 (2), 67–85.

Yang, Z., Chen, C., Lin, Y., Wang, D., Li, H., Xu, W., 2019. Effect of spatial enhancement technology on input through the keyboard in virtual reality environment. Appl. Ergon. 78, 164–175.

Zecevic, A., Miller, D., Harburn, K., 2000. An evaluation of the ergonomics of three computer keyboards. Ergonomics 43 (1), 55–72.

Zhang, M.R., Wobbrock, J.O., 2019. Beyond the input stream: making text entry evaluations more flexible with transcription sequences. In: Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, pp. 831–842.