

Controller-based Text-input Techniques for Virtual Reality: An Empirical Comparison

Costas Boletsis¹, Stian Kongsvik²

¹ SINTEF Digital, Forskningsveien 1, Oslo, Norway

² University of Oslo, Gaustadalléen 23B, Oslo, Norway

Corresponding author: Costas Boletsis, konstantinos.boletsis@sintef.no

Keywords: Controller – Human-computer interaction – Text input – User experience – Virtual reality

Abstract

Existing consumer VR systems support text input using handheld controllers in combination with virtual keyboards and many designers have attempted to build on these widely used techniques. However, information on current and well-established VR text-input techniques is lacking. In this work, we conduct a comparative empirical evaluation of four controller-based VR text-input techniques, namely, raycasting, drum-like keyboard, head-directed input, and split keyboard. We focus on their text-entry rate and accuracy, usability, and user experience. Twenty-two participants evaluated the techniques by completing a typing session, answering usability and user-experience questionnaires, and participating in a semi-structured interview. The drum-like keyboard and the raycasting techniques stood out, achieving good usability scores, positive experiential feedback, satisfactory text-entry rates, and moderate error rates that can be reduced in future studies. The specific documented usability and experiential characteristics of the techniques are presented and discussed herein.

1. Introduction

Since the early days of virtual reality (VR), various text-input techniques have been developed and studied to achieve seamless and user-friendly typing in virtual environments. Prior works have investigated many interaction methods for typing in VR, such as wearable gloves, specialised controllers, head and gaze direction, pen and tablet keyboards, virtual keyboards, touchscreen keyboards, augmented virtuality keyboards, speech-to-text, and hand and finger gestures (Lepouras, 2018; Lee and Kim, 2017; Grubert et al., 2018; Yu et al., 2017; McGill et al., 2015; Bowman, Rhoton, and Pinho, 2002).

Recently, VR has undergone a major hardware-driven revival, ushering in what has been characterised as the 'new era of virtual reality' (Smith, 2017; Olszewski et al., 2016; Boletsis, 2017). The introduction of the Oculus Rift Development Kit 1 in 2013 is considered a significant milestone for VR, indicating when the VR revival took place and when VR became accessible, up-to-date, and relevant again (Olszewski et al., 2016; Hilfert and König, 2016; Boletsis, Cedergren, and Kongsvik, 2017; Giuseppe and Wiederhold, 2015). The low acquisition cost of VR hardware has transformed VR into a popular technology that is widely accessible to researchers, designers, developers, and regular consumers. Moreover, the quality of virtual environments has improved rapidly, and they now offer realistic graphics and full immersion (Kim, Darakjian, and Finley, 2017; Moreira, de Oliveira, and Tori, 2016; Reinert et al., 2016). From a human-computer interaction (HCI) perspective, the technological revival of VR has led to the development of new and updated interaction tools for text input in VR, which have influenced the resulting user experiences (UX) and research in the domain (Boletsis, Cedergren, and Kongsvik, 2017).



Existing consumer VR systems, such as HTC Vive, Oculus Rift, and Samsung Gear VR, support text input using handheld controllers in combination with virtual keyboards (Grubert et al., 2018; Lee and Kim, 2017), and attempts to build on these widely used techniques are ongoing (Lee and Kim, 2017; Oberhauser and Lecon, 2017; George et al., 2017). However, the field lacks information about current and well-established VR text-input techniques. In this work, we conduct a comparative, empirical evaluation of four controller-based VR text-input techniques. This knowledge could help identify the interactions and experiential strengths and weaknesses of these widely used VR text-input techniques and guide the design of future VR systems (Kongsvik, 2018).

2. Related work

Since the introduction of the latest consumer VR systems, certain studies have utilized their interaction qualities to implement and evaluate various types of VR text-input techniques.

The integration of physical desktop keyboards with VR settings has attracted the attention of researchers. Walker et al. (2017) employed an orthogonal approach to examine the use of a completely visually occluded keyboard for typing in VR. The mean text-entry rates of their participants were 41.2–43.7 words-per-minute (WPM), with mean character error rates of 8.4%–11.8%. These character error rates were reduced to approximately 2.6%–4.0% through auto-correction of the typing input by using a decoder. McGill et al. (2015) investigated the process of typing on a desktop keyboard in augmented virtuality. Specifically, they compared a full keyboard view in reality with VR no-keyboard view, partial view and full blending conditions. They reported mean VR text-entry rates of 23.6, 38.5, and 36.6 WPM with mean total error rates (ER) of 30.86%, 9.2%, and 10.41%, respectively, under the three VR-related conditions. In addition, their results indicated that providing a view of the keyboard (VR partial view or full blending) positively influences typing performance. Under the same premise, Lin et al. (2017) examined conditions similar to those of McGill et al. (2015). They reported mean text-entry rates of 24.3–28.1 WPM and mean total ER of 20%–28%.

Grubert et al. (2018) investigated the performance of two desktop keyboards and two touchscreen keyboards for VR text entry. The mean text-entry rates achieved with the two desktop keyboard interfaces were 26.3 WPM and 25.5 WPM (and character error rates were 2.1% and 2.4%), separately. The mean text-entry rates achieved with the two touchscreen keyboard interfaces were 11.6 WPM and 8.8 WPM (character error rates: 2.7% and 3.6%), separately. The study of Grubert et al. (2018) confirmed that touchscreen keyboards were significantly slower than desktop keyboards, and novice users were able to retain approximately 60% of their typing speed on a desktop keyboard and about 40%–45% of their typing speed on a touchscreen keyboard. Head-based text entry in VR has been investigated as well. Gugenheimer et al. (2016) presented FaceTouch, an interaction concept in which head-mounted touchscreens are used to enable typing on the backside of head-mounted displays (HMD). In an informal user study with three experts, a text-entry rate of approximately 10 WPM was achieved with FaceTouch. Yu et al. (2017) studied a combination of head-based text entry with tapping (TapType), dwelling (DwellType), and gestures (GestureType). Users subjectively felt that all three techniques were easy to learn. The mean text-entry rates achieved with them were 15.58 WPM, 10.59 WPM, and 19.04 WPM, respectively, and the corresponding total ER were 2.02%, 3.69%, and 4.21%. A second study focused on the GestureType interface while improving the gesture-word recognition algorithm. A higher text-entry rate was achieved this time (24.73 WPM), but the total ER was higher (5.82%) as well.

Moreover, the original glove-based and controller-based techniques have been examined. Whitmire et al. (2017) presented and evaluated DigiTouch, a reconfigurable glove-based input device that enables thumb-to-finger touch interaction by sensing touch position and pressure continuously. In a series of 10 sessions, a

text-entry rate of 16 WPM (total ER: 16.65%) was achieved with DigiTouch in the last session. Lee and Kim (2017) presented a controller-based QWERTY-like touch-typing interface called Vitty, and they examined its usability for text input in VR compared to the conventional raycasting technique. Despite reported implementation issues, Vitty exhibited usability comparable to that of the raycasting technique, but its text-entry rate and accuracy were not examined.

Finally, in a previous study, we examined the VR drum-like keyboard with a focus on its text-entry rate and accuracy, usability, and UX (Boletsis and Kongsvik, 2019). The interface achieved a good usability score on the System Usability Scale (SUS), positive experiential feedback for its entertaining and immersive qualities, satisfactory text-entry rate (24.61 WPM), and moderate total ER (7.2%).

Most of the existing empirical studies pertaining to VR text-input techniques have focused on presenting newly constructed, original VR locomotion techniques and evaluating their text-entry performance. However, the HCI field of VR text input would benefit from an examination of existing under-researched VR text-input techniques, such as the controller-based ones, as a point of inspiration for new designs of VR text-input techniques. Moreover, an investigation of the experiential qualities of VR text-input techniques can help evaluate them comprehensively, which has not been done in most of the aforementioned studies. Exploratory empirical studies that investigate the performance and experiential characteristics of emerging and well-established VR text-input techniques can address these issues.

3. Controller-based VR text-input techniques

As the first step towards their empirical comparison, four controller-based VR text-input techniques were identified based on the authors' examination of consumer VR applications and the related literature (Yu et al., 2017; Grubert et al., 2018; Lee and Kim, 2017; Whitmire et al., 2017). All the selected techniques employ controllers for entering characters selected from virtual keyboards.

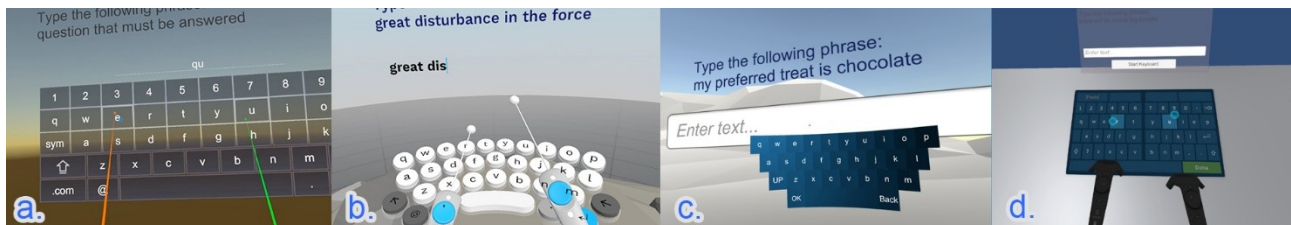


Figure 1: Controller-based VR text-input techniques evaluated and implemented in this study: a. Raycasting, b. Drum-like keyboard, c. Head-directed input, and d. Split keyboard (video demonstration: <https://boletsis.net/vrtext/>).

Raycasting: One of the most popular and conventional ways of text input in a VR setting is the 'aim and shoot' style, in which a hand-held controller is used to cast a virtual ray and select a particular key, and the final confirmation is made using a controller button (Lee and Kim, 2017). The two-handed ray casting technique requires a user to use both hands and hold two controllers for casting two rays (Fig. 1a).

Drum-like keyboard: The technique uses a drum set metaphor (Boletsis and Kongsvik, 2019). The controllers are used as sticks which – through downward movements – 'press' the keys of a virtual keyboard (Fig. 1b). The drum-like VR keyboard was presented as a prototype by Google Daydream Labs (Doronichev, 2016), and it was recreated in the context of open-source projects by Oculus' Jonathan Ravasz (Ravasz, 2017) and Normal VR company (Weisel, 2017).

Head-directed input: The user controls a pointer on a virtual keyboard by means of head rotation (Yu et al., 2017). The user selects a specific key and confirms it by pressing a button on the controller (Fig. 1c). The technique has dual functionality: head-based operation for key selection and controller-based operation for final choice confirmation.

Split keyboard: The technique employs a virtual keyboard split into two parts, one assigned to each controller. Thus, the user can type using both hands (Whitmire et al., 2017). In this implementation, key selection is made through the touch-sensitive trackpad of the Vive controller, and the final confirmation is made by pressing the trackpad button (Fig. 1d).

4. Evaluation study

A comparative study of the four aforementioned VR text-input techniques was conducted, with a focus on the text-entry rate and accuracy, usability, and UX of the techniques. The most widely used methodology for evaluating text-input interfaces involves presenting participants with preselected text phrases that they then enter using the text-input interface, and performance data are collected in the process (MacKenzie and Soukoreff, 2003). These phrases are usually retrieved randomly from a phrase set, such as the established MacKenzie and Soukoreff phrase set (MacKenzie and Soukoreff, 2003). In this study, this methodology was utilised, and text phrases were selected randomly from the MacKenzie and Soukoreff phrase set. The selected phrases are listed in Table 1 (mean phrase length: 29.7 characters, SD: 0.95, range: 28–31).

Table 1: Phrase set used in this study.

Phrase	Phrase length (characters)
- my preferred treat is chocolate	31
- question that must be answered	30
- there will be some fog tonight	30
- physics and chemistry are hard	30
- we are subjects and must obey	29
- great disturbance in the force	30
- wear a crown with many jewels	29
- my bank account is overdrawn	28
- movie about a nutty professor	29
- the king sends you to the tower	31

4.1. Interface & apparatus

All interfaces were developed on the Unity 3D game engine¹ and were deployed on the HTC Vive VR headset², and the Vive controllers were used. No haptic or vibratory feedback was implemented for keystrokes. Furthermore, no auto-completion or auto-correction functionalities were implemented to enable comparison with previous related works and to capture the baseline performance of the interface. A C# script was executed to generate a log file with various measurements (e.g. timings and keystrokes), and the values recorded in the log file were used to calculate the text-entry rate and accuracy of all techniques.

¹ <https://www.unity3d.com>

² <https://www.vive.com>

The raycasting technique was implemented using the Unity plugin Keyboard VR by Weelco Inc.³, and the drum-like VR keyboard was implemented using the open-source code of Punchkeyboard by Jonathan Ravasz⁴. The head-directed input technique was developed using the Unity plugin Curved VR Keyboard by Handcrafted VR⁵, and the split keyboard was developed using the OpenVR SDK by Valve Software⁶. All techniques featured a similar VR typing environment, with the VR keyboard in front of the user, typing entry box above the keyboard, and requested-phrase box above that (Figure 1).

4.2. Measures

4.2.1. Performance metrics

The dependent performance metrics used in this evaluation for examining the text-entry rate and accuracy were WPM and total ER.

Words-per-minute is perhaps the most widely reported empirical measure of text-entry performance (Wobbrock, 2007; Arif and Stuerzlinger, 2009). Since around 1905, a 'word' has commonly been regarded as five characters, including spaces (Yamada, 1980). The WPM measure does not consider the number of keystrokes or gestures made during entry; it considers only the length of the resulting transcribed string and the time required to produce it (Wobbrock, 2007). Thus, the formula for computing WPM is as follows (Wobbrock, 2007; Arif and Stuerzlinger, 2009):

$$WPM = \frac{|T|-1}{S} \times 60 \times \frac{1}{5}$$

Where T is the final transcribed string (phrase) entered by the subject, and $|T|$ is the length of this string. T may contain letters, numbers, punctuation, spaces, and other printable characters, but it may not contain backspaces. Thus, T does not capture the process of text entry but only the result of text entry (Wobbrock, 2007). The S term is seconds, and it is measured from the entry of the first character to the entry of the last character, which means that the entry of the first character is never timed; hence, '- 1' is included in the phrase length (Wobbrock, 2007; Arif and Stuerzlinger, 2009). The '60' denotes seconds per minute, and '1/5' denotes words per character.

Total ER is a unified metric that combines the effect of accuracy during and after text entry (Arif and Stuerzlinger, 2009; Soukoreff and MacKenzie, 2003). This metric measures the ratio of the total number of incorrect and corrected characters to the total number of correct, incorrect, and corrected characters (Soukoreff and MacKenzie, 2003):

$$Total\ ER = \frac{INF + IF}{C + INF + IF} \times 100$$

Where C denotes *correct* keystrokes, which are alphanumeric keystrokes that are not erroneous; INF denotes *incorrect and not fixed* keystrokes, which are errors that go unnoticed and appear in the transcribed text; and IF denotes *incorrect but fixed* keystrokes, which are erroneous keystrokes in the input stream that are later corrected (Soukoreff and MacKenzie, 2003).

In this evaluation, the *recommended* error correction condition was utilised, a condition that is frequently used in text-input evaluations, because it encourages normal user behaviour for correcting typing errors (Arif

³ <https://assetstore.unity.com/packages/tools/input-management/keyboard-vr-83143>

⁴ <https://github.com/rjth/Punchkeyboard>

⁵ <https://github.com/HandcraftedLabs/CurvedVRKeyboard>

⁶ <https://github.com/ValveSoftware/openvr>

and Stuerzlinger, 2009, 2010). Under this condition, participants can correct typing errors as soon as they identify them.

4.2.2. Questionnaires & interviews

Demographic data were collected in the initial stage of the study. These data included age, sex, and frequency of VR use ('never', 'rarely', 'frequently', and 'everyday').

For measuring usability, the 10-item *System Usability Scale (SUS)* questionnaire (Brooke, 2013) was used. This instrument allows usability practitioners and researchers to measure the subjective usability of products and services. Specifically, it is a 10-item questionnaire that can be administered quickly and easily, and it returns scores ranging from 0 to 100. Moreover, SUS scores can be translated into adjective ratings, such as 'worst imaginable', 'poor', 'OK', 'good', 'excellent', 'best imaginable', as well as into grade scales ranging from A to F (Bangor, Kortum, and Miller, 2009). The SUS has been demonstrated to be a reliable and valid instrument, robust with a small number of participants. In addition, it has the distinct advantage of being technology agnostic, meaning it can be used to evaluate a wide range of hardware and software systems (Brooke, 2013, 1996; Tullis and Stetson, 2004; Kortum and Acemyan, 2013).

User experience was measured using the *Game Experience Questionnaire (GEQ)* (IJsselsteijn, De Kort, and Poels, 2013), which has been used in several domains (such as gaming, augmented reality, and location-based services) because of its ability to cover a wide range of experiential factors with good reliability (Lee and Kim, 2017; Nacke, Grimshaw, and Lindley, 2010; Nacke and Lindley, 2008a,b; Lee et al., 2012). The use of GEQ has been established in the VR domain in several studies around such topics as navigation and locomotion in virtual environments (Meijer, Geudeke, and Van den Broek, 2009; Nabiyouni and Bowman, 2015), haptic interaction in VR (Ahmed et al., 2016), VR learning (Apostolellis and Bowman, 2014), cyberpsychology (Toet, van Welie, and Houtkamp, 2009), and VR gaming (Schild, LaViola, and Masuch, 2012). In this study, the dimensions of Competence, Sensory and Imaginative Immersion, Flow, Tension, Challenge, Negative Affect, Positive Affect, Returning to Reality, and Tiredness were selected from the In-Game and Post-Game versions of the GEQ based on the user instructions of the questionnaire (IJsselsteijn, De Kort, and Poels, 2013). This was done because it was necessary to probe the users' feelings and thoughts while typing and after they had stopped typing. The questionnaire asked the user to indicate how he or she felt during and after the session based on 19 statements (e.g. 'I forgot everything around me') on a five-point intensity scale ranging from 0 ('not at all') to 4 ('extremely').

The semi-structured interviews collected the participants' comments. The participants were asked about what they liked and did not like about the evaluated VR text-input techniques and the reasons thereof. The interviewer was able to follow up on the participants' comments until each topic was covered.

4.3. Participants

The participants were recruited from the authors' institutions. The recruited participants had to be physically able to use VR technology, and previous experience with VR was not a prerequisite. The participants were made aware of the potential risk of motion-sickness and the fact that they could opt out of the study at any time. All participants provided informed consent to participate in the study.

4.4. Procedure

The comparative study followed a within-subject design. First, the participants were presented with an introduction to the study, and they provided their informed consent. The participants then filled out the

demographic and VR-experience questionnaires. Then, the experimenters presented the first VR text-input technique to the participants, and they were given some trial time to familiarise themselves with the technique. Thereafter, the formal task commenced, and the participants were tasked with typing the 10 phrases listed in Table 1 as quickly and accurately as possible. The phrases were shown to the participants one at a time and were kept visible throughout the typing task. When the task was completed, the SUS and GEQ questionnaires were administered. A short break followed. The same procedure was followed for the remaining VR text-input techniques. After evaluation of the fourth technique, the semi-structured interview took place. The testing order of the VR text-input techniques was randomised.

4.5. Statistical Analysis

All data were analysed using the Statistical Package for Social Sciences (SPSS) version 25. The significance level was set to $p < 0.05$. Descriptive analysis was performed to depict the demographic data of the participants and to analyse the GEQ and SUS values. The non-parametric Friedman test was used to detect differences between the performance of the techniques based on the GEQ and SUS values. Repeated measures ANOVA was performed to compare the means of WPM and total ER of the four controller-based VR text-input techniques. The interview data were transcribed and subsequently analysed using open and axial coding, where the core concepts, themes, and ideas were identified. Two researchers coded the data independently, and the interrater reliability was assessed.

5. Results

5.1. Demographics

Twenty-two participants ($N = 22$, mean age: 25.77, SD: 5.28, male/female: 14/8) evaluated the four VR text-input techniques. Five participants had never experienced VR before, nine participants had experienced VR rarely, seven participants had experienced VR frequently, and one participant was experiencing VR daily. Among the 17 participants who had experienced VR previously, two had used HMD devices (e.g. Oculus Rift, HTC Vive, and PlayStation VR) and mobile VR headsets (e.g. Samsung Gear VR and Google Cardboard), 13 had used only HMD devices, and two participants had used only mobile VR headsets. All participants completed the sessions successfully.

5.2. Text-entry Rate and Accuracy

Tables 2 and 3 summarise the text-entry rates and accuracy results of the four controller-based VR text-input techniques. The repeated measures ANOVA indicated statistically significant differences among the mean WPM values, $F(2.169, 45.545) = 167.01$, $p < 0.001$, and the mean total ER values, $F(3, 63) = 4.794$, $p = 0.005$.

Table 2: Words-per-minute (WPM) performance of the evaluated controller-based VR text input techniques.

VR text input technique	WPM Mean (SD)	WPM Range
Raycasting	16.65 (3.28)	12.14 – 23.33
Drum-like keyboard	21.01 (3.7)	13.43 – 29.81
Head-directed input	10.83 (1.84)	7.52 – 14.85
Split keyboard	10.17 (2.39)	6.63 – 15.23

Table 3: Total error rates (total ER) of the evaluated controller-based VR text-input techniques.

VR text input technique	Total ER Mean (SD)	Total ER Range
Raycasting	11.05% (6.03%)	3.88% – 30.93%
Drum-like keyboard	12.11% (6.53%)	1.98% – 25.94%
Head-directed input	10.15% (3.74%)	4.19% – 19.73%
Split keyboard	8.11% (4.96%)	3.26% – 20.16%

5.3. SUS

The results of the SUS survey conducted herein are summarized in Table 4. These results were obtained based on the adjective ratings described by Brooke (1996, 2013).

The results of the Friedman test indicated statistically significant differences among the SUS values of the four techniques, $X^2(3) = 31.764$, $p < 0.001$.

Table 4: Descriptive statistics and ratings of the four evaluated VR text-input techniques, as obtained using the System Usability Scale scores.

VR text input technique	SUS Mean (SD)	SUS Range	SUS Rating
Raycasting	81.7 (13.85)	47.5 – 100	Good
Drum-like keyboard	85.34 (12.66)	52.5 – 100	Excellent
Head-directed input	66.7 (11.91)	47.5 – 85	OK
Split keyboard	66.59 (18.14)	27.5 – 95	OK

5.4. GEQ

Table 5 displays the mean values obtained using the GEQ questionnaire. As stated before, the values range from 0 ('not at all') to 4 ('extremely'). The results of the Friedman test indicate statistically significant differences in all GEQ dimensions, except for Returning to Reality, $X^2(3) = 0.241$, $p = 0.971$.

Table 5: Mean GEQ values (with standard deviation in parentheses) and the Friedman test results across the nine experiential dimensions of the evaluated controller-based VR text-input techniques.

GEQ dimensions	GEQ Mean (SD)				X^2 (p-value)
	Raycasting	Drum-like keyboard	Head-directed input	Split keyboard	
Competence	2.55 (0.84)	2.8 (0.87)	2.2 (0.7)	2.3 (0.85)	8.036 (p=0.045)
Immersion	2.41 (0.91)	3.11 (0.8)	2.16 (0.94)	2.34 (0.79)	31.043 (p<0.001)
Flow	1.59 (1.1)	2.27 (1.17)	1.7 (1.24)	2 (1.08)	8.695 (p=0.034)
Tension	0.45 (0.67)	0.3 (0.57)	1.25 (0.99)	0.91 (1)	18.304 (p<0.001)
Challenge	1.32 (1.02)	1.02 (1.02)	1.73 (1.14)	1.89 (1.06)	16.994 (p=0.001)
Negative Affect	0.86 (0.71)	0.3 (0.45)	1.15 (1)	0.75 (0.55)	28.914 (p<0.001)
Positive Affect	2.25 (0.99)	2.73 (0.88)	1.77 (0.91)	2.18 (0.76)	20.068 (p<0.001)
Returning to Reality	0.3 (0.48)	0.3 (0.46)	0.32 (0.49)	0.26 (0.34)	0.241 (p=0.971)
Tiredness	0.32 (0.48)	0.18 (0.33)	0.77 (0.84)	0.41 (0.7)	10.578 (p=0.014)

5.5. Interviews

Table 6 presents the comments of the participants collected from the interview sessions, together with the frequency of their occurrence. The participants comments are further characterised as positive and negative.

Table 6: Participants comments as collected from the interview sessions.

Raycasting		
Interview comment		Count
P	It is easy and fast to select and type characters with both hands.	13
N	It can be difficult to aim at the right character.	7
P	The integration of a keyboard and the game-like, shooting interaction metaphor adds a level of familiarity.	4
N	Both hands need to be moved to type fast, and that can be tiresome.	4
P	Choosing characters using laser beams and shooting can be fun.	4
Drum-like keyboard		
Interview comment		Count
P	It provides clear text-input feedback when the 'drumstick'/controller hits a key.	10
N	It is prone to error because hard hits can register a character twice.	8
P	It is an intuitive and fast technique because of the use of both hands and the fact that the drum interaction metaphor is familiar.	7
N	It can be tiresome because it requires active use of both hands.	5
P	'Playing the drums' to type is fun.	5
Head-directed input		
Interview comment		Count
N	Moving the head for typing can be very tiresome after a while.	11
N	It is difficult for the user to see what he or she is typing while typing it because any head movement will disrupt the user's key selection.	9
P	It is easy to press a key because the process relies on head movement, and it is clear where the user is gazing.	7
P	The technique is not tiresome in terms of hand use because the user is required to use just one hand and controller for confirming the key press.	2
Split keyboard		
Interview comment		Count
N	Assigning the keyboard's parts to specific controllers (e.g. left keyboard part is accessed only by the left controller) can affect users' typing style and freedom.	10
N	The use of the controller's touchpad makes it difficult to accurately select characters.	7
P	Using both hands and, specifically, thumbs adds flow and comfort to typing.	6
P	The controller's touchpad interface is familiar, because of the previous use of the laptop's touchpad.	3

*P: Positive comment, N: Negative comment.

6. Discussion

6.1. Utilised Measures

Methodologically, the use of the SUS and GEQ questionnaires and the semi-structured interviews, along with the text-input performance metrics, allowed for the discovery, verification, and documentation of significant experiential and interaction issues. The SUS and GEQ questionnaires provided a general overview of the usability and experiential performance of each method, whereas the interviews shed light on the specific interaction elements that the users liked and disliked. By combining all these measures, we managed not only to document how these techniques perform quantitatively but also why they perform the

way they do and how the users perceive their performance. However, the GEQ Returning to Reality dimension did not provide any significant comparative insights about the four controller-based VR text-input techniques.

6.2. Comparative Performance of Techniques

The drum-like keyboard exhibited superior performance relative to the other techniques. It yielded a high WPM rate (mean: 21.01, SD: 3.7) and high values on the GEQ dimensions of Competence, Immersion, Flow, and Positive Affect, and low values on the dimensions of Tension, Challenge, Negative Affect, and Tiredness. These positive GEQ values were further supported by the high SUS score realized with the technique (mean: 85.34, SD: 12.66) and the user interviews. The participants were satisfied with the clear text-input feedback and the familiarity of the drumming technique, and they enjoyed the playful drumming interaction metaphor. On the negative side, the drum-like keyboard technique yielded a higher total ER (mean: 12.11%, SD: 6.53%) than the other techniques, a fact that was confirmed from the interviews because participants found drumming to be error-prone because characters registered twice on hard hits. Moreover, some participants found the technique to be tiresome because of the active use of both hands for drumming. However, this was not a major complaint, as indicated by the GEQ Tiredness value.

The raycasting technique performed similarly to the drum-like keyboard. It yielded a lower but acceptable WPM value (mean: 16.65, SD: 3.28%) and a lower total ER (mean: 11.05%, SD: 6.03%). Raycasting performed well on the SUS scale (mean: 81.7, SD: 13.85). Many participants enjoyed the shooting technique and found it game-like and familiar. The technique's GEQ performance was similar to that of the drum-like keyboard on the GEQ dimensions of Competence, Immersion, Tension, and Positive Affect; yet, a few of the participants found it difficult to aim at the right key with this technique and tiresome to use both hands for 'shooting', which probably were the reasons for its lowest Flow score among all four techniques and second-ranked scores on Negative Affect.

The head-directed input and split-keyboard techniques had the lowest WPM values among the four techniques (mean: 10.83, SD: 1.84 and mean: 10.17, SD: 2.39, respectively) and the lowest error rates (mean: 10.15%, SD: 3.74% and mean: 8.11%, SD: 4.96%, respectively). Moreover, their perceived usability SUS scores were in the 'OK' rating range (mean: 66.7, SD: 11.91 and mean: 66.59, SD: 18.14, respectively), which are considered low scores. Their low usability and experiential performance is further supported by their low GEQ values (i.e. low values in positive dimensions and high values in negative dimensions) and interview remarks. Head-directed input exhibited the worst GEQ performance among the four techniques on the Competence, Immersion, Tension, Negative Affect, Positive Affect, and Tiredness dimensions. The performance of the split-keyboard technique was similar, but marginally superior, on the GEQ dimensions, except on the Challenge value, where it exhibited the worst performance among all techniques. The participants found the head-directed input to be tiresome and disruptive owing to constant head movement. However, they found that the confirmation action of the controller-based key selection was not tiresome for the hands. Moreover, the participants thought that the split keyboard limited their typing style and freedom by assigning parts of the keyboard to specific controllers, while their opinions on the controller touchpad method for selecting characters were divided. Seven participants found it difficult to use the touchpad for character selection, six participants claimed that the use of thumbs (on the touchpad) enhanced interaction comfort, and three participants found the touchpad to be familiar interaction-wise because of their previous experience of using the touchpad on their laptops.

6.3. General Observations

From a performance perspective, our evaluation of the four techniques confirmed the existence of differences in the text-entry rates, accuracy, and experiential elements of the drum-like keyboard and raycasting technique versus the head-directed input and the split-keyboard technique. The use of the drum-like VR keyboard and raycasting to type in VR resulted in promising mean text-entry rates. Especially, the drum-like keyboard could have achieved higher rates, closer to the mean values documented by Boletsis and Kongsvik (2019), if it were not for a few low-quality performances (as the SD and range statistics imply). These results suggest that the rates achieved with the drum-like keyboard and raycasting technique may be competitive against those of the other techniques discussed in Section 2, such as head-based (Yu et al., 2017; Gugenheimer et al., 2016), glove-based (Whitmire et al., 2017), and touchscreen-keyboard (Grubert et al., 2018) techniques. In addition, the techniques managed to perform similarly to several implementations of VR-integrated physical keyboards (Lin et al., 2017; McGill et al., 2015; Grubert et al., 2018). Naturally, some implementations of physical keyboards for VR text input, such as that of Walker et al. (2017), can achieve significantly superior rates, but their different use contexts should be highlighted in comparison. Physical keyboards can facilitate VR text entry for users in static, probably sitting, positions and office tasks, while the drum-like VR keyboard and the raycasting technique are used in various mobility and position settings (e.g. gaming), as well as to perform casual VR tasks (e.g. browsing, short communications), where the controller is the main interaction device (Boletsis and Kongsvik, 2019).

In terms of the accuracy of text entry, all techniques performed moderately, compared to the total ER of other techniques, without using text auto-correction or auto-completion functionalities, as described in Section 2. An approach to address and improve the accuracy of text entry is discussed in the Study Limitations subsection.

From a UX point of view, the evaluation study showed that the main reasons for which user prefers a VR text-input technique may not only be how fast they can type or how many errors they make when using it but also the enjoyment, agency, and positive emotions they get out of it. Moreover, physical elements can affect UX. Tiredness is an important factor when evaluating controller-based techniques, and, as can be seen in the interview remarks, there is no unified perspective. Therefore, some users may find a two-handed technique to be comfortable and fast because of the use of both hands while other users may find the same technique tiresome for the same reason.

6.4. Study Limitations

When analysing the evaluation results of the controller-based VR text-input techniques, additional factors should be considered.

The techniques examined herein followed a 'stripped' implementation, that is, without text auto-correction or auto-completion functionalities, and they were evaluated in only one session because of the exploratory nature of the study. Based on related literature (Grubert et al., 2018; Walker et al., 2017; Whitmire et al., 2017; Yu et al., 2017), a possible hypothesis for future research is that the text-entry rate and accuracy of these techniques can be improved by *i*) implementing decoders for text auto-correction and auto-completion and by *ii*) enabling users to complete several typing sessions so that they become more familiar with the interfaces. Moreover, a multi-session methodology can influence the WPM and total ER metrics, potentially resulting in superior performance. It can also influence several GEQ experiential dimensions, such as Tiredness, potentially resulting in higher values because of the interfaces' active physical interaction; Positive Affect, potentially resulting in lower values as the 'wow factor' wears off; and Flow and Competence, resulting in higher values because of participants' additional familiarity with the techniques.

6.5. Study Implications

Based on the results of this comparative empirical study of the four controller-based VR text-input techniques, we can list a few implications that can be useful to practitioners and researchers working in this domain.

First, evaluation of VR text-input techniques based solely on the text-entry and accuracy metrics may constitute a one-dimensional research approach. Examination of the techniques' experiential characteristics with a mixed-methods approach may shed more light on their overall performance, why users are or are not using them, and how they can be improved. All these elements are crucial for investigating the topic in a deeper fashion and advancing the HCI field.

Moreover, users should be given typing freedom and choices because 'one size does not fit all'. All the examined controller-based VR text-input techniques are similar in terms of their interface characteristics. Their co-existence and simultaneous inclusion in a specific VR task context that utilises controllers would be the optimal approach for the users to try and decide which technique they prefer. Our study showed that users may have completely opposite interaction experiences for the same reasons. Therefore, a single, optimal VR text-input technique may not be a realistic goal, unless it consists of several similar interfaces that facilitate various interaction metaphors that are interchangeable on a per-task and per-user basis.

Finally, the field of VR text input could benefit from exploratory comparative studies that analyse existing systems and shape the design of future systems. Based on the findings of this study, the drum-like keyboard is a promising VR text-input technique, and several VR applications can benefit from its implementation and integration. Nevertheless, an optimised key registration motion is necessary for this technique to reduce error rates. Along similar lines, raycasting would benefit from improved key selection and aiming, for example, by zooming in on the selected character. The split keyboard can be improved by allowing both controllers to access the entire VR keyboard. The head-directed input technique proved to be challenging and tiresome, but its implementation and interaction qualities can be further researched and adjusted based on the task at hand (e.g. adding an extra level of difficulty in VR games, VR typing for users with physical disabilities, etc.).

7. Conclusion

In this study, four controller-based VR text-input techniques were evaluated empirically: raycasting, drum-like VR keyboard, head-directed input, and split keyboard. In addition to text-entry rate and accuracy, the study managed to capture the experiential qualities of these techniques and the reasons that shape them through usability and UX questionnaires and semi-structured interviews. The drum-like keyboard and the raycasting technique stood out, achieving good usability scores, positive experiential feedback, satisfactory text-entry rates, and moderate error rates that can be further reduced in the future. Researchers and practitioners in the domain can benefit from the methodological aspects of this study, as well as from the discovered usability and experiential issues that can be addressed in future designs. In the future, we will examine the integration of text auto-correction and auto-completion functionalities and their effects on the text-input metrics and experiential qualities of the techniques. Moreover, a multi-session experimental design with a larger sample size will be implemented.

8. Acknowledgements

This research is funded by the Norwegian Research Council through the Centre for Service Innovation.

9. References

- Ahmed, I.; Harjunen, V.; Jacucci, G.; Hoggan, E.; Ravaja, N.; and Spape', M. M. 2016. Reach out and touch me: Effects of four distinct haptic technologies on affective touch in virtual reality. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, 341–348. ACM.
- Apostolellis, P., and Bowman, D. A. 2014. Evaluating the effects of orchestrated, game-based learning in virtual environments for informal education. In *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology*, 4. ACM.
- Arif, A. S., and Stuerzlinger, W. 2009. Analysis of text entry performance metrics. In *Proceedings of IEEE Toronto International Conference on Science and Technology for Humanity*, 100–105. IEEE.
- Arif, A. S., and Stuerzlinger, W. 2010. Predicting the cost of error correction in character-based text entry technologies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 5–14. ACM.
- Bangor, A.; Kortum, P.; and Miller, J. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies* 4(3):114–123.
- Boletsis, C., and Kongsvik, S. 2019. Text input in virtual reality: A preliminary evaluation of the drum-like vr keyboard. *Technologies* 7(2):31:1–31:10.
- Boletsis, C.; Cedergren, J. E.; and Kongsvik, S. 2017. HCI research in virtual reality: A discussion of problem-solving. In *Proceedings of International Conference on Interfaces and Human Computer Interaction*, 263–267. IADIS.
- Boletsis, C. 2017. The new era of virtual reality locomotion: A systematic literature review of techniques and a proposed typology. *Multimodal Technologies and Interaction* 1(4):24:1–24:17.
- Bowman, D. A.; Rhoton, C. J.; and Pinho, M. S. 2002. Text input techniques for immersive virtual environments: An empirical comparison. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2154–2158. SAGE.
- Brooke, J. 1996. SUS - A quick and dirty usability scale. In Jordan, P.; Thomas, B.; Weerdmeester, B.; and McClelland, I., eds., *Usability Evaluation in Industry*. Taylor & Francis. 189–194.
- Brooke, J. 2013. SUS: A Retrospective. *Journal of Usability Studies* 8(2):29–40.
- Doronichev, A. 2016. Daydream Labs: exploring and sharing VR's possibilities. <https://blog.google/products/google-vr/daydream-labs-exploring-and-sharing-vrs/>.
- George, C.; Khamis, M.; von Zezschwitz, E.; Schmidt, H.; Burger, M.; Alt, F.; and Hußmann, H. 2017. Seamless and secure vr: Adapting and evaluating established authentication systems for virtual reality. In *Proceedings of Workshop on Usable Security*, 1–12. Internet Society.
- Giuseppe, R., and Wiederhold, B. K. 2015. The new dawn of virtual reality in health care: medical simulation and experiential interface. *Studies in Health Technology & Informatics* 219:3–6.
- Grubert, J.; Witzani, L.; Ofek, E.; Pahud, M.; Kranz, M.; and Kristensson, P. O. 2018. Text entry in immersive head-mounted display-based virtual reality using standard keyboards. In *Proceedings of IEEE Conference on Virtual Reality and 3D User Interfaces*, 159–166. IEEE.
- Gugenheimer, J.; Dobbstein, D.; Winkler, C.; Haas, G.; and Rukzio, E. 2016. Face-touch: Enabling touch interaction in display fixed UIs for mobile virtual reality. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, 49–60. ACM.
- Hilfert, T., and König, M. 2016. Low-cost virtual reality environment for engineering and construction. *Visualization in Engineering* 4(1):2.
- IJsselstein, W.; De Kort, Y.; and Poels, K. 2013. *The Game Experience Questionnaire*. Technische Universiteit Eindhoven.
- Kim, A.; Darakjian, N.; and Finley, J. M. 2017. Walking in fully immersive virtual environments: An evaluation of potential adverse effects in older adults and individuals with Parkinson's disease. *Journal of neuroengineering and rehabilitation* 14(1):16.
- Kongsvik, S. 2018. Text input techniques in virtual reality environments – an empirical comparison. Master's thesis, University of Oslo.
- Kortum, P., and Acemyan, C. Z. 2013. How low can you go?: is the system usability scale range restricted? *Journal of Usability Studies* 9(1):14–24.
- Lee, Y., and Kim, G. J. 2017. Vitty: Virtual touch typing interface with added finger buttons. In *Proceedings of International Conference on Virtual, Augmented and Mixed Reality*, 111–119. Springer.
- Lee, G. A.; Dunser, A.; Kim, S.; and Billinghamurst, M. 2012. CityViewAR: A mobile outdoor AR application

- for city visualization. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*, 57–64. IEEE.
- Lepouras, G. 2018. Comparing methods for numerical input in immersive virtual environments. *Virtual Reality* 22(1):63–77.
- Lin, J.-W.; Han, P.-H.; Lee, J.-Y.; Chen, Y.-S.; Chang, T.-W.; Chen, K.-W.; and Hung, Y.-P. 2017. Visualizing the keyboard in virtual reality for enhancing immersive experience. In *ACM SIGGRAPH 2017 Posters*, 35. ACM.
- MacKenzie, I. S., and Soukoreff, R. W. 2003. Phrase sets for evaluating text entry techniques. In *Proceedings of CHI '03 Extended Abstracts on Human Factors in Computing Systems*, 754–755. ACM.
- McGill, M.; Boland, D.; Murray-Smith, R.; and Brewster, S. 2015. A dose of reality: Overcoming usability challenges in VR head-mounted displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2143–2152. ACM.
- Meijer, F.; Geudeke, B. L.; and Van den Broek, E. L. 2009. Navigating through virtual environments: Visual realism improves spatial cognition. *CyberPsychology & Behavior* 12(5):517–521.
- Moreira, P.; de Oliveira, E. C.; and Tori, R. 2016. Impact of immersive technology applied in computer graphics learning. In *Proceedings of the Brazilian Symposium on Computers in Education*, 410–419. SBIE.
- Nabiyouni, M., and Bowman, D. A. 2015. An evaluation of the effects of hyper-natural components of interaction fidelity on locomotion performance in virtual reality. In *Proceedings of the 25th International Conference on Artificial Reality and Telexistence and 20th Eurographics Symposium on Virtual Environments*, 167–174. Eurographics Association.
- Nacke, L., and Lindley, C. A. 2008a. Boredom, immersion, flow: A pilot study investigating player experience. In *Proceedings of the IADIS International Conference Gaming 2008: Design for engaging experience and social interaction*, 1–5. IADIS Press.
- Nacke, L., and Lindley, C. A. 2008b. Flow and immersion in first-person shooters: measuring the player's gameplay experience. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, 81–88. ACM.
- Nacke, L. E.; Grimshaw, M. N.; and Lindley, C. A. 2010. More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interacting with Computers* 22(5):336–343.
- Oberhauser, R., and Lecon, C. 2017. Immersed in software structures: A virtual reality approach. In *Proceedings of International Conference on Advances in Computer-Human Interactions*, 181–186. IARIA.
- Olszewski, K.; Lim, J. J.; Saito, S.; and Li, H. 2016. High-fidelity facial and speech animation for VR HMDs. *ACM Transactions on Graphics* 35(6):221.
- Ravasz, J. 2017. Keyboard Input for Virtual Reality. <https://uxdesign.cc/keyboard-input-for-virtual-reality-d551a29c53e9>.
- Reinert, B.; Kopf, J.; Ritschel, T.; Cuervo, E.; Chu, D.; and Seidel, H.-P. 2016. Proxy-guided image-based rendering for mobile devices. *Computer Graphics Forum* 35(7):353–362.
- Schild, J.; LaViola, J.; and Masuch, M. 2012. Understanding user experience in stereoscopic 3D games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 89–98. ACM.
- Smith, P. A. 2017. Using commercial virtual reality games to prototype serious games and applications. In *Proceedings of International Conference on Virtual, Augmented and Mixed Reality*, 359–368. Springer.
- Soukoreff, R. W., and MacKenzie, I. S. 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*, 113–120. ACM.
- Toet, A.; van Welie, M.; and Houtkamp, J. 2009. Is a dark virtual environment scary? *CyberPsychology & Behavior* 12(4):363–371.
- Tullis, T. S., and Stetson, J. N. 2004. A comparison of questionnaires for assessing website usability. In *Proceedings of the Usability Professional Association Conference*, 1–12. Usability Professional Association.
- Walker, J.; Li, B.; Vertanen, K.; and Kuhl, S. 2017. Efficient typing on a visually occluded physical keyboard. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 5457–5461. ACM.
- Weisel, M. 2017. An open-source keyboard to make your own. <http://www.norm.alvr.com/blog/an-open-source-keyboard-to-make-your-own/>.

- Whitmire, E.; Jain, M.; Jain, D.; Nelson, G.; Karkar, R.; Patel, S.; and Goel, M. 2017. Digitouch: Reconfigurable thumb-to-finger input and text entry on head-mounted displays. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 113:1–113:21. ACM.
- Wobbrock, J. O. 2007. Measures of text entry performance. In Card, S.; Grudin, J.; and Nielsen, J., eds., *Text entry systems: Mobility, accessibility, universality*. Morgan Kaufmann. 47–74.
- Yamada, H. 1980. A historical study of typewriters and typing methods, from the position of planning Japanese parallels. *Journal of Information Processing* 2:175– 202.
- Yu, C.; Gu, Y.; Yang, Z.; Yi, X.; Luo, H.; and Shi, Y. 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*, 4479–4488. ACM.