

Investigation of the User Experience of Text Entry in Virtual Reality Applications

Practical Report

Steffen Kautz, BSc User Experience Design (Hons), 2021

[4981 Words]

Abstract

Several studies have investigated the user-experience of different text input methods for virtual reality (VR) applications. Of those achievable with standard VR controllers, ambidextrous pointer input has been found to be most effective. One other study aimed to find the optimal vergence for VR user interfaces. The experiment conducted for this report aimed to investigate the effect of different positions of the virtual keyboard on text input efficiency and comfort of users. The results indicate a preferable vergence at 2.7m distance, but also show the need to improve the method and perform further tests to obtain more reliable results.

Contents

Abstract _____	2
Introduction _____	4
History of keyboards	
History of VR	
Recent developments	
Current challenges	
Chapter 1 – Text Input _____	5
Text input methods	
Means to increase speed and accuracy	
Chapter 2 – Methodology _____	7
Task design	
Prototype design	
Measures	
Participants	
Procedure	
Chapter 3 – Results and Discussion _____	9
Results	
Discussion	
Conclusion _____	10
Appendix _____	11
Bibliography _____	15

Introduction

History of Keyboards

The first application of a keyboard as we know it today was in the form of typewriters in 1866. It was invented by the Americans C. L. Sholes and C. Glidden while “working on a machine for consecutively numbering rail tickets, bank notes and the pages of books” (Noyes, 1983). They developed the typewriter including the first version of the now ubiquitous QWERTY character layout. Since then, keyboard have been adopted as a staple of human computer interaction. They have been an essential asset to the spread of text-based electronic communication and continue to be embedded in any kind of communication device from personal computers to mobile phones. There is a range of primary human computer interfaces across different devices. Physical keyboards are supplemented with computer mice, or substituted with controllers that have much fewer buttons, and since the invention of capacitive touch screens they have become a common mode of input, too. With those different means of controlling devices, virtual keyboards have replaced physical ones, and with their virtualisation, new interaction methods have been derived. Smartphone keyboards needed to be improved to make up for the user only being able to control them with two fingers at a time. As a solution to this challenge they suggest a list of words the user is likely to type, and swipe keyboards improve speed by removing the need to lift one’s finger off the keyboard or be entirely accurate for that matter (Kung, Hsie, Smith, 2021).

History and uses of Virtual Reality

Since its conception in the early twentieth century, the notion of complete escapism into a simulated environment has been a topic of scientific interest. An early concept of Virtual Reality (VR) dates back to 1935, when Stanley G. Weinbaum published his story “Pygmalion’s Spectacles”. The earliest practical applications of VR technology was in the entertainment sector. For example, Morton Heilig used it in the 1950s to create immersive film displays with stimulation of a variety of senses, with a machine he called the “Sensorama” (Craig, Sherman, Will, 2009). Helmet- or head-mounted displays (HMDs), which now are practically synonymous with VR, were first deployed in military flight simulators developed by Thomas A. Furness for the United States Airforce (USAF) in 1966. These simulators were later developed further into the heads-up display systems used in fighter pilots’ helmets including augmented reality targeting systems, which are in active use today (Bye, 2015).

In the early 1990s, SEGA attempted to establish consumer VR headsets in the video game market. Due to hardware limitations at the time their value did not hold up to consumer expectations and they turned out not to be economically feasible (ThrillSeeker, 2020).

Recent Developments

Technological development is still very far off the omni-sensual immersion envisioned by Weinbaum, but recently, VR facilitated by HMDs has re-emerged as a consumer technology. Due to advancements in small high-resolution displays, the quality of HMDs increased, while entry cost to VR experiences has dropped (ThrillSeeker, 2020). Different modes of input have emerged, like gaze interactions, hand tracking, and motion-tracked controllers (Kauhanen et al., 2017). VR is now applied in a range of fields, from professional to private applications: there are action-games like Beat Saber and Superhot VR, players for videos and movies supporting 3D content like Bigscreen, and creative applications revolving around three-dimensional painting and sculpting like Google’s Tilt Brush and Gravity Sketch.

In the professional field, XR (virtual, augmented and mixed reality) applications have found early adoption in industry branches involved in product design and engineering, like the architecture and automotive industries. Beyond design, VR can also be found in health care applications, where surgeons use it to prepare for operations, or are even assisted by augmented reality overlays during surgery, as demonstrated in Cydar’s CydarEV tool used in the preparation and execution of endovascular surgery (Cydar, 2021). VR is also still commonplace in flight simulator training for pilots, with full-flight simulators being purpose-built for different types of aircraft and used in mandatory preparation to obtain type ratings (CAE, 2021).

Current Challenges

As development of better and more economical hardware progresses, VR technology becomes accessible to more consumers. Still the main focus of VR applications remains on active and passive entertainment. There are some early pushes into adopting XR headsets for productivity applications, like Immersed Inc.’s virtual desktop application and diverse companies utilising Microsoft’s HoloLens platform, mainly for medical training and architectural visualisation. Still, some challenges remain, before work in VR becomes feasible. One of these challenges is text input. Currently the standard mode of interaction with VR applications are tracked controllers, while few use hand-tracking. Both methods have ground to cover regarding tracking accuracy, but even with perfect tracking neither of them would be preferable over a physical keyboard for longform text input. For tasks requiring short inputs, however, dropping controllers and keeping a tracked keyboard around is not feasible. This is why virtual keyboards are still an important aspect of VR applications and their further development is an essential part of improving the user experience.

Chapter 1 – Text Input

Methods

Both controllers and hand tracking are inaccurate to different extents, and especially input using VR controllers is an unfamiliar mode of input to most users. When it comes to text input, complexity is high and accuracy is paramount for a good user experience. In order to enable users to maintain input speed and accuracy, there is room to develop solutions to aid data input. There are various pathways from which this can be approached, with focus ranging from hardware devices to software implementations and from algorithms to interaction modes (Elmgren, 2017; Speicher et al., 2018; Dudley, 2019):

- 1) Speech recognition
- 2) Handwriting recognition (Fig.2)
- 3) Physical keyboard
 - a. Tracked full-size keyboard
 - b. Pinch keyboard (Fig. 3)
- 4) Virtual keyboard (Fig. 1)
 - a. Head pointing
 - b. Controller pointing
 - c. Controller tapping
 - d. Freehand
 - e. Continuous cursor
 - f. Discrete cursor

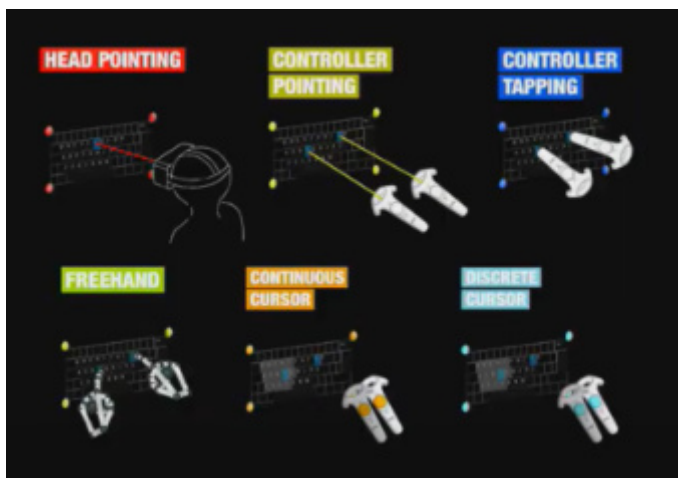


Figure 1 Input modalities for virtual keyboards (ACM SIGCHI, 2018)

Speech recognition is problematic regarding accuracy and privacy. Commands can easily be overheard and signal integrity can be compromised by external noise outside of the user's control (Speicher et al., 2018). It also performs very poorly with inputs of rare words, phonetically identical words especially out of context, and non-standard phrases often used in computing, such as strings containing case sensitive or special characters. Based on this, speech recognition is only viable as a complimentary input method.

Dudley (2019) suggests the use of probabilistic interfaces for text input via hand tracking, using computational

models to infer intention and balance noisy input with the likelihood of particular inputs. This includes taking into account the direction and speed at which the cursor was moving before the selection was made, the distance of the selected point from the centre of each key in the proximity, and the relative chance of either character following in the string typed so far based on a dictionary and potentially profiles of users' word usage. Weighing all these factors against each other enables the software to make very accurate predictions and reduce errors in text input.

Research conducted by Elmgren (2017) suggests that, while text input via handwriting is perceived as fun and engaging, virtual keyboards outperform it significantly in terms of efficiency. Key issues with this method are the lack of friction in combination with humans jittering naturally. When writing by hand, we are used to rest our hands on the surface we are writing on as a point of reference. Without this reference, natural tremors mean input becomes very noisy and hard to parse by the program. It also goes against the original reason for the invention of keyboards: to find a method of text generation that is faster than handwriting.



Figure 2 Experimental implementation of handwriting recognition input in VR (Elmgren, 2017)

A tracked, full-size, physical keyboard outperforms any virtual solution in terms of input speed, due to its familiarity and tactile feedback. The downside is, though, that it limits users in VR applications, as it inhibits mobility in the physical space and confines users to a working place or requires them to move a physical item around with them that cannot be attached to their body or the VR set, introducing a factor of inconvenience. Tracked physical keyboards currently available, as implemented in the ImmersedVR app for Oculus systems, also require the keyboard to be connected to a computer which is accessed via the remote desktop application (Cas and Chary VR, 2020). ImmersedVR also does not feature avatar hands, since current Oculus devices, like most other HMDs, do not have the tracking facilities to accurately map a virtual skeleton to the user's hands and transpose them into VR. A study conducted by Knierim et al. (2018) shows that, while avatar hands only have little impact on typing performance, they do enhance presence significantly and aid in reducing

workload imposed on typists during text input. In conclusion, physical keyboards excel at text input, but given the dominant modes of interaction with VR applications at this time, switching between controllers and a keyboard for text input is not worthwhile for short-form input.

Pinch keyboards are held in one hand and operated by pinching down on the buttons. They typically have 12 keys in 3 rows. A full range of inputs can be achieved with button combinations referred to as chords. This type of input has a high learning curve and is generally slow. Another drawback is that it requires a device that is only viable for this one type of operation, making the skills needed to use it not only hard to obtain, but also non-transferable.

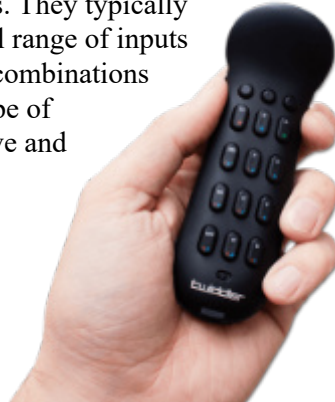


Figure 3 Twiddler3 pinch keyboard (Tekgear, 2016)

On virtual keyboards, a study conducted by Speicher et al. (2018) shows that controller pointing is the best option in all measured aspects bar physical demand. The methods that they compared were head pointing, continuous and discrete cursor input, freehand typing, controller tapping, and controller pointing. When head pointing the user points their head at the key corresponding to the character they wish to type. Holding their gaze still over the letter for a specific amount of time types the character. This mode of input requires a reasonable amount of effort, moving and stabilising the head with the added weight of the HMD. Furthermore it is quite slow, with a relatively lengthy amount of time needed to ensure no accidental inputs are made. Continuous and discrete cursor text selection are techniques developed for entering text on game consoles and televisions using controllers or remote controls. Two keys are highlighted with selectors, each of which are controlled with a corresponding thumb-stick on the controllers. When an entry button is pressed, the selected character is entered. For the discrete cursor method, each cursor is confined to either half of the keyboard, whereas with the continuous cursor method each cursor can be navigated across the entire keyboard. This text entry method is proven to work with little physical strain, but text input speed is negatively impacted by the amount of time needed to select characters at potentially long distances across the virtual keyboard. This is partially mitigated by the use of two cursors instead of just one as common in previous applications of this technique. Tracking two cursors at the same time and remembering which one corresponds to which hand places a higher mental load on the user's working memory though, which increases the effort needed to type. When entering text with the freehand method, the user's

hands are tracked and their movements matched to the virtual keyboard. This works reasonably well, apart from the physical effort and strain produced from the lack of resistance and haptics. Without a point to rest their wrists, users have to exert a significant amount of effort to stabilise their hands and maintain accuracy without haptic feedback. Thus the typing experience is not comparable enough to that of a physical keyboard, making it harder to learn and more exhausting to use. Lastly with current consumer HMDs, none offer accurate enough hand tracking to facilitate this method of input on their own, instead requiring third party sensors that have to be worn in addition to the HMD, creating a higher barrier to entry and further increasing physical strain. Controller tapping works similarly, but uses the controller position tracking already in place, instead of tracking the user's hands. This alleviates the barrier to entry and the effort of dropping the controllers and picking them back up, but it also inhibits input speed, by reducing the interacting points from up to ten fingers to two arms, which now have to make larger movements resulting in more physical exhaustion. Controller pointing, which was found to be the most effective and one of the least straining methods, casts rays from both controllers, which are used to select keys. The only input methods resulting in less physical strain were the discrete and continuous cursor methods, which result in far slower typing speeds. From the current perspective on short-form VR text entry, this appears to be the most efficient mode of input. Depending on the type of input required, it can be enhanced using the methods described by Dudley (2019) among others.

The study conducted by Speicher et al. (2018) does not measure the impact of different positions of the virtual keyboard on the user experience. This area will be more closely examined in this experiment, where I aim to develop a testing method, apply it, and finally assess the results and method.

Chapter 2 – Methodology

Task Design

For safety and comfort reasons the test will be performed with subjects sitting and using the VR set. Some testers are physically disabled and would experience pain standing for extended periods of time. In order to keep results comparable between subjects, all participants will be accommodated in the same way.

The chosen input method is likely to be complimentary to a physical keyboard for long-form text input. Thus the focus will be on short forms of input:

- a familiar string (the test subject's name)
- an intelligible string with special characters (file name or e-mail address)
- a random string including special characters (password)

Each of these test phrases will be repeated once per pass for each keyboard position

Prototype Design

The prototype was developed using A-Frame, a JavaScript (JS) based framework for WebVR developed by MozVR (2015). This framework integrates into HTML code, describing three-dimensional objects which can be arranged to a scene with a range of properties. Entities can be marked with identifiers and be given properties that link them to logic programmed in JS in the same file. A-Frames can be run in almost any browser, including the Oculus Quest's native browser, which was used for this experiment.

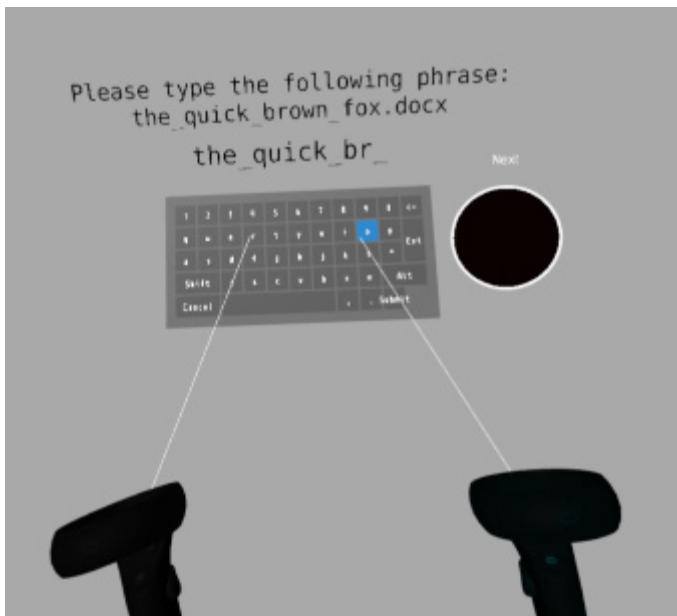


Figure 4 Screenshot of the A-Frames prototype used for testing

To compliment it the "A-Frame Keyboard" library published by Wanderer OÜ was used as an asset to bring a virtual keyboard into the VR environment. It encompasses a square-grid qwerty keyboard with limited special characters on the default layout and a wider selection upon pressing the "Alt" key. During testing

users were given time to familiarise themselves with this non-standard layout before starting the tasks. This choice of keyboard was influenced by the temporal constraints on the development of the prototype. For more accurate testing results an entirely standard keyboard layout is preferable.

Reading comfort is essential for users, so keyboard and text display distances for the tests reference those established by Dingler, Kunze and Outram (2018). In their research they determined the median closest comfortable distance at 0.89m, the lower quartile most comfortable distance at 1.3m, the median most comfortable distance at 2.7m, the upper quartile most comfortable distance at 8.6m, and the median farthest comfortable distance at 9.8m. To determine if closer or further vergence is beneficial to text input, this experiment examines median, as well as upper and lower quartile most comfortable viewing distances. Regarding keyboard height there is a comparison between elbow-height with the input field close to the keyboard, elbow-height with the input field at eye level, and the keyboard at eye-level with the input field close to the keyboard. This is aimed at determining the importance of comfort in relation to input location.

The keyboard is set up so the distance is determined at the start of the application, with users being stationary. Keeping keyboard distance consistent with users' head movements is not favourable as it can provoke motion sickness (Balabanian, Legkov, 2016; Parker, Prothero, 2003; Reason, 1975).

Measures

To assess performance of different keyboard placements I will employ three objective measures to determine input speed and accuracy. These measures have been adapted from Knierim et al. (2018, p.4-6) under consideration of the limitations of my prototype.

Objective Measures

Speed

Characters per Minute (CPM)

Characters per minute are calculated by measuring the time in seconds taken to complete a test phrase (t). The number of characters in the input is then divided by the time taken and the result multiplied with 60, to interpolate words typed to the span of one minute (Bowman, Rhoton and Pinho, 2002). Since the experiment focuses on short-form inputs, characters per minute are a more appropriate measurement in order to compare beyond the scope of this experiment than words per minute.

Accuracy

Number of Corrections

The number of characters deleted during input of a test phrase. (Knierim et al., 2018) This measure records characters accidentally or incorrectly entered.

Error Rate

The error rate measures differences between the original test phrase and the text input by the tester. To measure it, the Minimum String Distance (MSD) or Levenshtein Distance (Speicher et al., 2018; Soukoreff, MacKenzie, 2001) between the test phrase (P) and the entered text (T) is determined and divided by the character count of the longer string. Multiplying the result of this operation with 100 yields the error rate in percent (Knierim et al., 2018; Levenshtein, 1966). Combining error rate and number of corrections will provide a picture of overall input accuracy.

Subjective Measures

Motion Sickness

To assess test participants' levels of motion sickness their reactions were recorded using the Motion Sickness Assessment Questionnaire (MSAQ) developed by Giaranos et al. (2001). This questionnaire is used to determine if participants experienced simulator sickness during the test, as this might impede accuracy of test results regarding the examined aspects of text input. The MSAQ measures motion sickness in 4 dimensions: gastrointestinal, central, peripheral, and sopite-related. The gastrointestinal dimension includes feelings of nausea, sickness, queasiness and the need to vomit. The central dimension of motion sickness relates to light headedness, dizziness and faint-like or spinning feelings. Feeling warm or cold, as well as sweat or clamminess, are part of the peripheral dimension. Lastly annoyance, irritation, drowsiness, fatigue, and feelings of unease belong to the sopite-related dimension.

Effort

In order to track the effort users experience with text input, the Task Load Questionnaire (NASA TLX) was adapted to record their perceived strain. The NASA TLX questionnaire covers a high level overview of effort, frustration, and demand, as well as participants' perceived performance, comparing the actuality against their expectations of the task. Measurements for demand are separated into physical, mental, and temporal demand. The goal is to find a balance between size of movements and concentration on fine motor skills. Since performance is assessed with objective measurements and frustration is outside of the scope of this experiment, these measures were omitted from the questionnaire. The TLX questionnaire was presented after every set of three phrases, to record comparable results for each position of the keyboard.

Subjects

Due to concerns regarding the ongoing Covid-19 pandemic, I decided to limit the pool of participants to persons I usually interact with. Thus my experiment draws on a narrow demographic and results may only be representative for users within these constraints. All test subjects were tested for Covid-19 before participating. In preliminary questioning it was ensured that none of the

test participants have epilepsy or are affected by conditions of the vestibular system. The test group consisted of 8 subjects between the ages of 18 and 24 years old. Of these 5 were male, 2 female, and 1 non-binary. All but one described themselves as experienced typists and comfortable with the QWERTY keyboard layout used in this test. In regard to experience with using VR applications, 6 stated they had used VR headsets a few times, and 2 stated they had no experience whatsoever. No experienced users of VR headsets were available for testing. 5 of the 8 participants wear prescription glasses and were asked to keep these on while using the headset, as the Oculus quest allows for spectacles to be worn underneath it. Typing accuracy measurements will be examined with particular care for 2 subjects who are dyslexic.

Setup and Procedure

All testing was performed in accordance with Norwich University of the Art's Research Ethics policy. The participants were asked preliminary questions regarding their demographic and to make sure nobody's health was at risk during the experiment. They were then explained what the experiment was meant to test, and how they should contribute to an accurate set of measurements. They were introduced to the VR headset and all best practices and safety precautions relating to its use. After the induction, they were asked to sit down, adjust the headset, and begin with the test. Upon completion of the test, they were asked to complete the motion sickness assessment questionnaire. Every test session was recorded for assessment, and after each session, the equipment was cleaned.

Problems with the setup

The prototype keyboard was not responsive enough to keep up with the input speed of some participants. This may have affected results to the point of being unable to measure differences in typing speed between sets, as users were limited by the restrictions of the prototype rather than by their perception of the keyboard. The cause for this limitation is likely to be the implementation of a complex object like a keyboard with a large number of interactive and dynamic elements in A-Frame, which is not optimised to handle this amount of complexity. For further testing it is advisable to use a more advanced prototype with a better integrated keyboard solution to guarantee input speed is not inhibited by the software's responsiveness. A good platform to attempt this in would be the Unity game engine, which would also not restrict the prototype to a web-based application run in the browser and possibly enable automatic data gathering in the prototype itself.

Chapter 3 – Results and Discussion

	Sec. 1	Sec. 2	Sec. 3	Sec. 4	Sec. 5	Sec. 6
CPM	33.98	38.33	36.20	37.84	41.15	41.46
Corrections	1.08	0.63	0.75	0.50	0.46	0.63
Levenshtein	0.25	0.13	0.17	0.13	0.04	0.00

Figure 5 Average results across all tests

Results

Test results were gathered manually on questionnaires and by reviewing the recorded footage of all test runs. The results of the motion sickness questionnaire show that there were no significant cases of motion sickness during testing. Results are therefore not affected by participants' discomfort. None of the measured dimensions showed any particular notability over the others. The overall highest average score was 3.13 out of 9 for disorientation. This is likely due to the low level of experience with VR applications in the sample group. One subject reported notably higher scores than others but reported this was due to their general health on that day rather than the experience during testing. They felt well enough to participate.

Values on the abridged NASA Task Load Index questionnaire varied greatly between test subjects. This is likely due to the questionnaire being based on subjective, self-reported scores. The results reported on a scale of 1 to 21 for mental load, physical load, temporal effort, and overall effort varied from a minimum of 1 to a maximum of 18.

Measuring input speed, the CPM for the name input Task 1 of each section varied greatly. The reason for this variance is likely the inconsistency in the length of input, ranging from 3 to 15 characters between subjects, as well as some test participants choosing not to capitalise their names. The average CPM for Task 1 ranges from 46.88 to 55.64, while Task 2 ranges from 30.16 to 39.91, and Task 3 ranges from 23.68 to 33.33. This decrease in input speed between the tasks was expected, as the progression of phrases goes from very familiar via intelligible to random. Furthermore, Task 2 required scanning for special characters, and Task 3 required more complex keyboard state changes between lower and upper case, as well as special characters.

Regarding accuracy measurements, higher scores mean lower accuracy. Both dyslexic participants showed exceptionally high scores on Task 2 in the first section. Since their Task 2 accuracy scores were much lower on the rest of the sections, they will be disregarded for section 1 assuming that they were due to subjects' adjustment to the test phrase. Overall, average accuracy scores ranged from 0 to 1, with no discernible difference between the test sections.

Discussion

The NASA Task Load Index questionnaire results are inconclusive. They seem to reduce slightly over the course of the experiment, but it is unclear whether this is

due to reducing load from each section to the next, or if it is the effect of the test subjects getting used to the task.

The values do however indicate that the task load is higher for wearers of prescription glasses and significantly higher for the dyslexic in relation to their respective counterparts. Since the results of the Task Load Index questionnaire are self-reported, it is however difficult to accurately compare them between participants, especially with a sample as small as the one in this experiment.

The objective measures show consistent improvement over the course of the experiment, which suggests that the subjects' performance was affected more by learning than by the different test setups (Fig. 6-7). Accuracy is increasing alongside input speed, which suggests an increase in overall skill. This could have been remedied by rotating the order of test sections for each subject, to introduce a regular amount of noise into the measurements. Proceeding in that manner would distribute the learning progress of each test subject and reduced its effect on the measured values.

Despite this effect it seems plausible to assume that a medium vergence is preferable to a far one, which in turn is preferable to a close focal distance. This interpretation relies on the slight decrease in speed and accuracy between sections 2 and 3, as well as just accuracy between sections 5 and 6. Overall it seems that medium to far vergence is the best performing distance, although additional research is needed to confirm this.

When comparing CPM between subjects for Tasks 2 and 3, the average difference between the fastest and slowest results are 28.94 CPM and 19.79 CPM respectively. Comparing input speed between sections shows an average difference of 9.75 and 9.66 CPM. This suggests that keyboard placement and vergence have a lesser effect on objective text input performance than the individual users' experience and dexterity.

The improvement of results within sections 1 to 3 and sections 4 to 6 is smaller than that between the two groups of sections, which indicates that having the keyboard at eye-height may be beneficial to the user experience.

Conclusion

Although the sample used for this test is small, there are indications in line with the findings of Dingler, Kunze, and Outram (2018) that medium vergence is preferable, not just for users' comfort, but also to aid text input speed and accuracy.

Furthermore, distinct improvements to the testing process have been identified, that can improve the reliability and accuracy of further tests, be obfuscating the learning effect in results and removing limitations that the prototype used in this experiment imposes on users' performance.

Improvements

During the testing and evaluation process two main areas for improvement have been identified:

- 1) Prototype design
 - a. Test phrase selection
 - b. Data gathering
 - c. Prototype performance
- 2) Testing process
 - a. Test order
 - b. Tester training
 - c. Tester frustration
 - d. Sample size

Prototype Design

The results during testing of familiar strings using the test subjects' names produced clearly distinct results from the other two tests. Therefore, taking into account the variations due to inconsistent input lengths, it is necessary to develop an exercise with similarly familiar inputs, while using a common test phrase.

Integrating all non-preliminary questionnaires into the testing programme and automating data gathering, not only for data gathered from questionnaires but also the input tasks, would make testing on larger scales than the limited scope of this experiment feasible.

A more stable prototype with closer control of testing parameters and support for depth independent sizing of UI elements would make testing more reliable, as well as enable more flexible tuning of the testing parameters with lower effort.

Testing Process

The aforementioned learning effect due to users' increasing familiarity with the test phrases and keyboard layout can be remedied by randomising test order for each participant and normalising results for linear improvement. It can also be reduced by training subjects with a more extensive warm-up exercise. For this it is important not to prime testers for just one of the tested setups, but rather alternate between all of them evenly. This together with the use of a completely standard keyboard layout can also make it easier for participants to find the required special characters more easily.

Another factor of inaccuracy of this experiment was test subjects' frustration with repetitive tasks. Due to typing the same phrases over and over again, participants became increasingly impatient and annoyed with the test. This frustration can be counteracted by reducing the number of tests run, or separating them into multiple sessions.

Lastly, increasing the sample size and gathering test subjects from a wider demographic range will ensure that the results are more representative of the general population.

It is recommended to pursue further research, taking into account these improvements to the method.

Appendix

Demographic data and NASA-TLX

Question	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7	Subject 8	average	max delta	MAX	MIN	No Dyslexia	Dyslexia	No Glasses	Glasses	No Dyslexia	Dyslexia	No Glasses	Glasses	
Epilepsy	No	No	No	No	No	No	No	No													
Vestibular Condition	No	No	No	No	No	No	No	No													
Dyslexia	No	No	No	No	No	No	Yes	Yes													
Prescription Glasses	Yes	Yes	No	No	Yes	Yes	No	Yes													
Age	23.00	20.00	22.00	20.00	21.00	18.00	23.00	24.00													
Gender	Female	Male	Male	Non-binary	Male	Female	Male	Male													
VR Experience	A little	A little	A little	A little	A little	None	None	A little													
Typing Expertise	Intermediate	Intermediate	Intermediate	Intermediate	Intermediate	Intermediate	Novice	Intermediate													
Section 1									average	max delta	MAX	MIN									
Mental Load	12.00	8.00	3.00	6.00	1.00	2.00	16.00	15.00	7.88	15.00	16.00	1.00	5.33	15.50	8.33	7.60	5.33	15.50	8.33	7.60	
Physical Load	2.00	6.00	1.00	4.00	1.00	2.00	5.00	13.00	4.25	12.00	13.00	1.00	2.67	9.00	3.33	4.80	2.67	9.00	3.33	4.80	
Temporal Load	9.00	3.00	3.00	2.00	1.00	2.00	5.00	12.00	4.63	11.00	12.00	1.00	3.33	8.50	3.33	5.40	3.33	8.50	3.33	5.40	
Overall Effort	7.00	9.00	7.00	4.00	1.00	3.00	12.00	14.00	7.13	13.00	14.00	1.00	5.17	13.00	7.67	6.80	5.17	13.00	7.67	6.80	
Section 2																					
Mental Load	9.00	9.00	2.00	2.00	1.00	2.00	10.00	11.00	5.75	10.00	11.00	1.00	4.17	10.50	4.67	6.40	4.17	10.50	4.67	6.40	
Physical Load	3.00	2.00	1.00	3.00	1.00	2.00	3.00	10.00	3.13	9.00	10.00	1.00	2.00	6.50	2.33	3.60	2.00	6.50	2.33	3.60	
Temporal Load	8.00	6.00	3.00	1.00	1.00	2.00	3.00	9.00	4.13	8.00	9.00	1.00	3.50	6.00	2.33	5.20	3.50	6.00	2.33	5.20	
Overall Effort	3.00	5.00	3.00	3.00	1.00	2.00	10.00	10.00	4.63	9.00	10.00	1.00	2.83	10.00	5.33	4.20	2.83	10.00	5.33	4.20	
Section 3																					
Mental Load	6.00	4.00	1.00	2.00	1.00	2.00	12.00	10.00	4.75	11.00	12.00	1.00	2.67	11.00	5.00	4.60	2.67	11.00	5.00	4.60	
Physical Load	3.00	2.00	3.00	2.00	1.00	2.00	5.00	9.00	3.38	8.00	9.00	1.00	2.17	7.00	3.33	3.40	2.17	7.00	3.33	3.40	
Temporal Load	3.00	3.00	2.00	1.00	1.00	2.00	3.00	10.00	3.13	9.00	10.00	1.00	2.00	6.50	2.00	3.80	2.00	6.50	2.00	3.80	
Overall Effort	5.00	4.00	3.00	4.00	1.00	2.00	15.00	8.00	5.25	14.00	15.00	1.00	3.17	11.50	7.33	4.00	3.17	11.50	7.33	4.00	
Section 4																					
Mental Load	2.00	7.00	3.00	3.00	1.00	1.00	10.00	16.00	5.38	15.00	16.00	1.00	2.83	13.00	5.33	5.40	2.83	13.00	5.33	5.40	
Physical Load	2.00	8.00	1.00	2.00	1.00	1.00	5.00	17.00	4.63	16.00	17.00	1.00	2.50	11.00	2.67	5.80	2.50	11.00	2.67	5.80	
Temporal Load	3.00	5.00	2.00	1.00	1.00	1.00	3.00	14.00	3.75	13.00	14.00	1.00	2.17	8.50	2.00	4.80	2.17	8.50	2.00	4.80	
Overall Effort	2.00	8.00	1.00	3.00	1.00	1.00	15.00	18.00	6.13	17.00	18.00	1.00	2.67	16.50	6.33	6.00	2.67	16.50	6.33	6.00	
Section 5																					
Mental Load	2.00	9.00	2.00	5.00	1.00	1.00	5.00	16.00	5.13	15.00	16.00	1.00	3.33	10.50	4.00	5.80	3.33	10.50	4.00	5.80	
Physical Load	3.00	10.00	2.00	8.00	1.00	1.00	3.00	17.00	5.63	16.00	17.00	1.00	4.17	10.00	4.33	6.40	4.17	10.00	4.33	6.40	
Temporal Load	1.00	5.00	2.00	1.00	1.00	1.00	3.00	13.00	3.38	12.00	13.00	1.00	1.83	8.00	2.00	4.20	1.83	8.00	2.00	4.20	
Overall Effort	6.00	9.00	2.00	9.00	1.00	1.00	5.00	15.00	6.00	14.00	15.00	1.00	4.67	10.00	5.33	6.40	4.67	10.00	5.33	6.40	
Section 6																					
Mental Load	4.00	4.00	2.00	3.00	1.00	1.00	5.00	12.00	4.00	11.00	12.00	1.00	2.50	8.50	3.33	4.40	2.50	8.50	3.33	4.40	
Physical Load	3.00	3.00	3.00	3.00	1.00	1.00	3.00	13.00	3.75	12.00	13.00	1.00	2.33	8.00	3.00	4.20	2.33	8.00	3.00	4.20	
Temporal Load	2.00	2.00	2.00	1.00	1.00	1.00	3.00	9.00	2.63	8.00	9.00	1.00	1.50	6.00	2.00	3.00	1.50	6.00	2.00	3.00	
Overall Effort	3.00	4.00	2.00	3.00	1.00	1.00	5.00	12.00	3.88	11.00	12.00	1.00	2.33	8.50	3.33	4.20	2.33	8.50	3.33	4.20	

Mental Load	Physical Load	Temporal Load	Overall Effort
7.88	4.25	4.63	7.13
5.75	3.13	4.13	4.63
4.75	3.38	3.13	5.25
5.38	4.63	3.75	6.13
5.13	5.63	3.38	6.00
4.00	3.75	2.63	3.88

Motion Sickness Questionnaire

Motion Sickness	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7	Subject 8	average	max delta	MAX	MIN	No Dyslexia	Dyslexia	No Glasses	Glasses
Sick to stomach	1.00	2.00	1.00	1.00	1.00	4.00	1.00	1.00	1.50	3.00	4.00	1.00	1.67	1.00	1.00	1.80
Faint-like	1.00	1.00	2.00	2.00	1.00	1.00	1.00	1.00	1.25	1.00	2.00	1.00	1.33	1.00	1.67	1.00
Annoyed/Irritated	3.00	3.00	2.00	1.00	1.00	3.00	3.00	4.00	2.50	3.00	4.00	1.00	2.17	3.50	2.00	2.80
Sweaty	1.00	4.00	1.00	2.00	3.00	1.00	3.00	3.00	2.25	3.00	4.00	1.00	2.00	3.00	2.00	2.40
Queasy	1.00	1.00	1.00	1.00	1.00	2.00	1.00	5.00	1.63	4.00	5.00	1.00	1.17	3.00	1.00	2.00
Lightheaded	2.00	1.00	3.00	4.00	1.00	1.00	1.00	5.00	2.25	4.00	5.00	1.00	2.00	3.00	2.67	2.00
Drowsy	2.00	2.00	1.00	1.00	1.00	1.00	1.00	6.00	1.88	5.00	6.00	1.00	1.33	3.50	1.00	2.40
Clammy/Cold Sweat	1.00	6.00	1.00	1.00	2.00	1.00	1.00	1.00	1.75	5.00	6.00	1.00	2.00	1.00	1.00	2.20
Disoriented	3.00	4.00	1.00	2.00	1.00	4.00	5.00	5.00	3.13	4.00	5.00	1.00	2.50	5.00	2.67	3.40
Tired/Fatigued	2.00	2.00	1.00	1.00	1.00	1.00	7.00	4.00	2.38	6.00	7.00	1.00	1.33	5.50	3.00	2.00
Nauseated	1.00	2.00	1.00	1.00	1.00	3.00	1.00	2.00	1.50	2.00	3.00	1.00	1.50	1.50	1.00	1.80
Hot/Warm	1.00	3.00	2.00	1.00	1.00	1.00	3.00	5.00	2.13	4.00	5.00	1.00	1.50	4.00	2.00	2.20
Dizzy	2.00	1.00	3.00	3.00	1.00	3.00	1.00	6.00	2.50	5.00	6.00	1.00	2.17	3.50	2.33	2.60
Spinning	1.00	1.00	1.00	1.00	1.00	1.00	1.00	3.00	1.25	2.00	3.00	1.00	1.00	2.00	1.00	1.40
May Vomit	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00
Uneasy	2.00	3.00	1.00	1.00	1.00	3.00	3.00	3.00	2.13	2.00	3.00	1.00	1.83	3.00	1.67	2.40

Averages for Speed and Accuracy

Averages		Section 1	Section 2	Section 3	Section 4	Section 5	Section 6		Difference
Task 1	Characters	7.00	7.00	7.00	7.00	7.00	7.00		
	Time Taken	9.00	8.25	10.75	9.00	8.75	8.75		2.50
	CPM	48.09	53.20	47.49	46.88	55.64	51.13		8.76
	Corrections	0.38	0.50	0.63	0.63	0.50	0.75		0.38
	Levenshtein	0.00	0.00	0.13	0.00	0.00	0.00		0.13
Task 2	Characters	24.00	24.00	24.00	24.00	24.00	24.00		
	Time Taken	52.50	43.50	45.13	42.00	42.00	37.38		15.13
	CPM	30.16	34.12	36.21	36.56	36.00	39.91		9.75
	Corrections	2.75	1.00	1.25	0.88	0.63	1.13		2.13
	Levenshtein	0.63	0.00	0.25	0.13	0.13	0.00		0.63
Task 3	Characters	12.00	12.00	12.00	12.00	12.00	12.00		
	Time Taken	32.38	26.50	30.38	26.50	24.38	22.38		10.00
	CPM	23.68	27.68	24.91	30.08	31.81	33.33		9.66
	Corrections	0.13	0.38	0.38	0.00	0.25	0.00		0.38
	Levenshtein	0.13	0.38	0.13	0.25	0.00	0.00		0.38
Combined									
	CPM	33.98	38.33	36.20	37.84	41.15	41.46		
	Corrections	1.08	0.63	0.75	0.50	0.46	0.63		
	Levenshtein	0.25	0.13	0.17	0.13	0.04	0.00		
Task 1	Characters	7.83	7.83	7.83	7.83	7.83	7.83		
	Time Taken	9.00	9.00	12.17	9.33	9.50	8.67		3.50
	CPM	54.13	54.26	49.82	51.56	58.19	58.18		8.36
	Corrections	0.17	0.50	0.67	0.50	0.50	0.67		0.50
	Levenshtein	0.00	0.00	0.00	0.00	0.00	0.00		0.00
Task 2	Characters	24.00	24.00	24.00	24.00	24.00	24.00		
	Time Taken	44.33	42.50	44.33	40.50	41.83	34.17		10.17
	CPM	33.80	35.16	33.60	38.28	36.34	42.52		8.91
	Corrections	0.33	0.67	0.83	0.67	0.67	1.00		0.67
	Levenshtein	0.67	0.00	0.17	0.17	0.17	0.00		0.67
Task 3	Characters	12.00	12.00	12.00	12.00	12.00	12.00		
	Time Taken	31.00	26.83	28.83	24.33	22.33	22.17		8.83
	CPM	24.62	27.49	26.06	31.73	33.88	33.68		9.26
	Corrections	0.17	0.50	0.17	0.00	0.00	0.00		0.50
	Levenshtein	0.00	0.17	0.00	0.17	0.00	0.00		0.17
Combined									
	CPM	37.51	38.97	36.50	40.52	42.80	44.79		
	Corrections	0.22	0.56	0.56	0.39	0.39	0.56		
	Levenshtein	0.22	0.06	0.06	0.11	0.06	0.00		

Bibliography

Alger, M. (2015) Vr interface design manifesto(2014). Available at: <https://www.youtube.com/watch?v=n3b8hZ5NV2E> (Accessed: 12 October 2021).

Balabanian, A. and Legkov, P. (2016) 005 - motion sickness in vr: adverse health problems in vr part i, Research VR Podcast - The Science & Design of Virtual Reality. Available at: <https://researchvr.podigee.io/5-researchvr-005> (Accessed: 12 October 2021).

Boas, Y. A. G. V. (2012) 'Overview of virtual reality technologies'. Available at: <https://www.semanticscholar.org/paper/Overview-of-Virtual-Reality-Technologies-Boas/4214cb09e29795f5363e5e3b545750dce027b668> (Accessed: 15 November 2021).

Bowman, D. A., Rhoton, C. J. and Pinho, M. S. (2002) 'Text input techniques for immersive virtual environments: an empirical comparison', Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 46(26), pp. 2154–2158. doi: 10.1177/154193120204602611.

Bye, K. (no date) 50 years of vr with tom furness: the super cockpit, virtual retinal display, hit lab, & virtual world society | voices of vr podcast. (Voices of VR). Available at: <https://voicesofvr.com/245-50-years-of-vr-with-tom-furness-the-super-cockpit-virtual-retinal-display-hit-lab-virtual-world-society/> (Accessed: 15 November 2021).

CAE (2021) Full-flight simulators. Available at: <https://www.cae.com/civil-aviation/aviation-simulation-equipment/training-equipment/full-flight-simulators/> (Accessed: 15 November 2021).

Cas and Chary VR (2020) You can now bring your real keyboard in vr & it's awesome! Available at: <https://www.youtube.com/watch?v=rwXGOABYAY> (Accessed: 12 October 2021).

Coomans, M. K. D. and Timmermans, H. J. P. (1997) Towards a taxonomy of virtual reality user interfaces. London: International Conference on Information Visualisation.

Craig, A. B., Sherman, W. R. and Will, J. D. (2009) Developing virtual reality applications: foundations of effective design. Burlington, MA : Oxford: Morgan Kaufmann ; Elsevier Science [distributor].

Cydar (2021) CydarEV, Cydar Medical. Available at: <https://www.cydarmedical.com/product> (Accessed: 15 November 2021).

Dingler, T., Kunze, K. and Outram, B. (2018) 'Vr reading uis: assessing text parameters for reading in vr', in Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18: CHI Conference on Human Factors in Computing Systems, Montreal QC Canada: ACM, pp. 1–6. doi: 10.1145/3170427.3188695.

Dudley, J. J. (2020) Probabilistic user interface design for virtual and augmented reality applications. Thesis. University of Cambridge. doi: 10.17863/CAM.51660.

Elmgren, R. (2017) 'Handwriting in vr as a text input method', undefined. Available at: <https://www.semanticscholar.org/paper/Handwriting-in-VR-as-a-Text-Input-Method-Elmgren/9ab3eaf1e25522b0982f97edc67b50dac4dbf5af> (Accessed: 12 October 2021).

Gianaros, P. J. et al. (2001) 'A Questionnaire for the Assessment of the Multiple Dimensions of Motion Sickness'. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2910410/pdf/nihms215311.pdf>.

Google (2017) Designing screen interfaces for vr(Google i/o '17). Available at: <https://www.youtube.com/watch?v=ES9jArHRFHQ> (Accessed: 12 October 2021).

How virtual reality became a reality (2020). Available at: <https://www.youtube.com/watch?v=UW8NpTA9oc> (Accessed: 12 October 2021).

- Kauhanen, O. et al. (2017) 'Assisting immersive virtual reality development with user experience design approach', in Proceedings of the 21st International Academic Mindtrek Conference. AcademicMindtrek'17: Annual Academic Mindtrek Conference, Tampere Finland: ACM, pp. 127–136. doi: 10.1145/3131085.3131126.
- Kharoub, H., Lataifeh, M. and Ahmed, N. (2019) '3d user interface design and usability for immersive vr', Applied Sciences, 9(22), p. 4861. doi: 10.3390/app9224861.
- Knierim, P. et al. (2018) 'Physical keyboards in virtual reality: analysis of typing performance and effects of avatar hands', in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18: CHI Conference on Human Factors in Computing Systems, Montreal QC Canada: ACM, pp. 1–9. doi: 10.1145/3173574.3173919.
- Kung, C.-H., Hsieh, T.-C. and Smith, S. (2021) 'Usability study of multiple vibrotactile feedback stimuli in an entire virtual keyboard input', Applied Ergonomics, 90, p. 103270. doi: 10.1016/j.apergo.2020.103270.
- Laugwitz, B., Held, T. and Schrepp, M. (2008) 'Construction and evaluation of a user experience questionnaire', in Holzinger, A. (ed.) HCI and Usability for Education and Work. Berlin, Heidelberg: Springer (Lecture Notes in Computer Science), pp. 63–76. doi: 10.1007/978-3-540-89350-9_6.
- Levenshtein, V. I. (1966) 'Binary codes capable of correcting deletions, insertions, and reversals', Soviet Physics-Doklady, February, pp. 707–710. Available at: <https://nymity.ch/sybilhunting/pdf/Levenshtein1966a.pdf>.
- MozVR (2015) Introducing a-frame: building blocks for webvr, Mozilla Mixed Reality Blog. Available at: <https://blog.mozvr.com/introducing-aframe/> (Accessed: 18 October 2021).
- Noyes, J. (1983) 'The QWERTY keyboard: a review', International Journal of Man-Machine Studies, 18(3), pp. 265–281. doi: 10.1016/S0020-7373(83)80010-8.
- Parker, D. and Prothero, J. (2003) 'A unified approach to presence and motion sickness', in Hettinger, L. and Haas, M. (eds) Virtual and Adaptive Environments. CRC Press, pp. 47–66. doi: 10.1201/9781410608888.ch3.
- Reason, J. D. (1975) Motion Sickness. London: Academic Press.
- Soukoreff, R. W. and 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. Seattle, Washington: Association for Computing Machinery (CHI EA '01), pp. 319–320. doi: 10.1145/634067.634256.
- Speicher, M. et al. (2018) 'Selection-based text entry in virtual reality', in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18: CHI Conference on Human Factors in Computing Systems, Montreal QC Canada: ACM, pp. 1–13. doi: 10.1145/3173574.3174221.
- Usuh, M. et al. (2000) 'Using presence questionnaires in reality', Presence: Teleoperators and Virtual Environments, 9(5), pp. 497–503. doi: 10.1162/105474600566989.
- WandererOU (2021) 'A-Frame keyboard readme'. GitHub. Available at: <https://github.com/WandererOU/aframe-keyboard/blob/master/README.md> (Accessed: 18 October 2021).