Made available by Hasselt University Library in https://documentserver.uhasselt.be

Exploring Alternative Text Input Modalities in Virtual Reality: A Comparative Study

Peer-reviewed author version

JANS, Mathias; CEYSSENS, Jeroen & LUYTEN, Kris (2024) Exploring Alternative Text Input Modalities in Virtual Reality: A Comparative Study. In: Spencer, SN (Ed.). 30Th ACM Symposium on Virtual reality Software and Technology, VRST 2024, ASSOC COMPUTING MACHINERY, p. 1 -2.

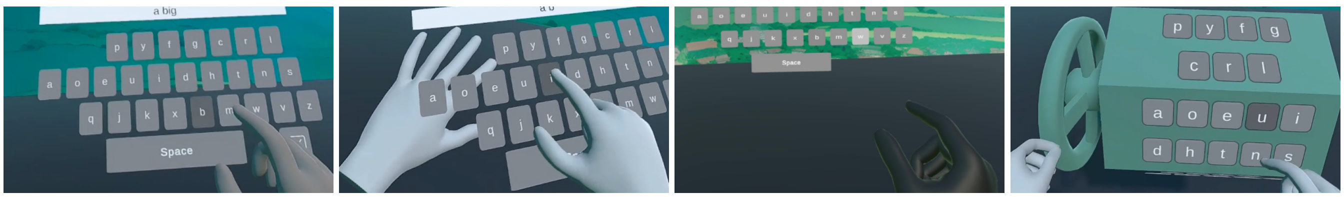
DOI: 10.1145/3641825.3689505
Handle: http://hdl.handle.net/1942/44877

Exploring Alternative Text Input Modalities in Virtual Reality: A Comparative Study

Mathias Jans
mathias.jans@uhasselt.be
UHasselt - Flanders Make
Expertise Centre for Digital Media
Diepenbeek, Belgium

Jeroen Ceyssens
jeroen.ceyssens@uhasselt.be
UHasselt - Flanders Make
Expertise Centre for Digital Media
Diepenbeek, Belgium

Kris Luyten
kris.luyten@uhasselt.be
UHasselt - Flanders Make
Expertise Centre for Digital Media
Diepenbeek, Belgium



(a) Air **(b) Hand** **(c) Eyes** **(d) Shape**

Figure 1: The four keyboard implementations we compared, all using the Dvorak layout: (a) a floating keyboard with touch input, (b) a keyboard attached on the back of the hand with touch input, (c) a floating keyboard with eye tracking and pinch input, and (d) a keyboard laid out over a rolling shape with touch input.

Abstract

Text input in Virtual Reality (VR) is crucial for communication, search, and productivity. We compared four keyboard designs for VR text entry, leveraging the flexibility and the tracking options of a 3D environment. We used the Dvorak layout to control for experience differences. The designs were: (a) a floating keyboard with touch input, (b) a keyboard attached on the back of the hand with touch input, (c) a floating keyboard with eye tracking and pinch input, and (d) a keyboard laid out over a rolling shape with touch input. Designs (b), (c), and (d) can move in 3D space, while design (a) is static. Design (d) had similar efficiency to design (a) but with better usability and lower Physical Demand. Design (b) led to higher Physical Demand, Effort, and Frustration. Design (c) had lower Physical Demand but higher Mental Demand, Effort, and error rates. Typing speeds averaged 6.51 WPM (1.24% error rate) for (a), 5.56 WPM (3.82% error rate) for (b), 5.33 WPM (1.43% error rate) for (c), and 6.70 WPM (1.64% error rate) for (d).

CCS Concepts

• **Human-centered computing** → **Text input**; **Virtual reality**.

Keywords

Text input, Virtual Reality, Interface design

1 Introduction

The ability to enter text in Virtual Reality (VR) is essential for many applications, such as communicating with others, entering search queries, and for productivity applications. Historically, VR text entry relied on controllers [3], but recently we see a shift to alternative methods such as hand and eye tracking [5] as the technology became more readily available in both professional and consumer devices. This study explores hand and eye tracking for VR text input using the Dvorak keyboard layout. This choice avoids biases from more common layouts such as QWERTY, allowing us to focus on the interaction techniques. The Dvorak layout is designed to group frequently used keys together to create a more efficient layout, rather than avoiding typewriter issues as with the original QWERTY design. By leveraging this unique key organization, we analyze how it can enhance VR text input [4].

2 Methodology

2.1 Keyboard Designs

The four keyboard designs from our study, shown in Figure 1, are:

- **(a) Air** Virtual keyboard floats statically in front of the user.
- **(b) Hand** Virtual keyboard floats above the back of the hand.
- **(c) Eyes** Virtual keyboard floats in front of the user, with input via eye tracking and a pinch gesture.
- **(d) Shape** A 3D shape with key rows mapped on its surfaces. The shape can be rotated using a wheel by pinching it. Every row/surface represents a row in Dvorak (middle is most common keys, top is less common, bottom is least common).

Designs (a), (b) and (d) use the index finger for key selection, while design (c) uses eye tracking and pinching the index finger and thumb for input. All designs used the Dvorak layout to standardize results. Design (a) was the baseline, (b) tested the effect of moving the keyboard closer, (c) tested reduced arm movement, and (d) explored the impact of focusing on a portion of the keys at a time.

2.2 Apparatus

We used a Meta Quest Pro head-mounted display (HMD) with built-in hand and eye tracking, connected to a laptop for tethered rendering. The laptop featured an AMD Ryzen 7 5800H CPU, 32 GB of RAM, and an NVIDIA GeForce RTX 3070 GPU. Designs were created in Unity 2021.3.19f1, and statistical analysis was conducted in R 4.0.4. For keyboards that used touch input, we utilized the "XRHands" and "OpenXR Plugin" packages. Design (c) employed the "Oculus XR Plugin" for eye tracking and pinch confirmation. When design (c) is started, calibration was done using the Quest Pro's default eye tracking calibration. Because there was a different package used for the hand tracking, the hand looks different.

2.3 Procedure

Participants sat on a static chair without armrests during VR tests and at a table for the questionnaires, using paper forms. After signing an informed consent, they filled out demographic questionnaires and received study instructions. They practiced the Dvorak layout on a physical keyboard by typing four sentences. In VR, they typed five sentences with each design, which were randomized using a balanced Latin Square. They were not required to correct mistakes but could use the backspace key to delete text from their current position up to and including the point of the mistake. After each design, they filled out questionnaires. Upon finishing all designs, they ranked the keyboards in a final questionnaire.

2.4 Measures

Typing speed was measured in *words per minute (WPM)*, while *word error rate (WER)* assessed errors in the final sentence and *keystrokes per character (KSPC)* evaluated mistakes during typing. Questionnaires included the *Raw NASA Task Load Index (RTLX)* [2] for stress and fatigue, and the *System Usability Scale (SUS)* [1] for overall usability. Participants also ranked the keyboards on ease of use, typing performance, annoyance, and pleasantness.

2.5 Participants

The study included sixteen participants (eleven male, five female, ages 18-34). Thirteen were right-handed and three left-handed. One person had never used VR before, seven barely used it, four used it occasionally, one used it frequently, and three used it very frequently. None had experience with the Dvorak layout.

3 Results

For detailed results, refer to Table 1 in Appendix A. We applied Repeated Measures ANOVA for normally distributed data and the Friedman Rank Sum test for non-normally distributed data, with pairwise Wilcoxon rank sum tests and Bonferroni correction ($\alpha = 0.05$) for post-hoc analysis. No significant effects of *handedness* or *VR experience* were found. Among the designs, *Shape* had the highest average typing speed (M=6.70 WPM), slightly ahead of *Air* (M=6.51 WPM) without a significant difference. Both were significantly faster than *Eyes* (M=5.33) and *Hands* (5.56 WPM). Typing speeds were lower than those reported by Schenkluhn [5], who achieved 11.67 WPM with gaze and pinch and 15.27 WPM with touch input, likely due to unfamiliarity with the Dvorak layout. The *WER* ranged from 1.24% to 3.82%, with no significant differences among designs, indicating effective error correction. However, *Eyes* had an error rate of 1.43%, lower than Schenkluhn's 6.08% for gaze and pinch. Our other designs' rates compared to his tap method's 2.02%. Our values are higher than ideal, likely due to limited error correction mechanisms. *Eyes* showed the highest *KSPC* (more mistakes), significantly different from other designs. Questionnaire results showed significance for *RTLX* Physical Demand (best with *Shape* and *Eyes*, third *Air*), Effort (best *Shape*, second *Air*), and Frustration (best *Shape* and *Eyes*). For the *SUS* results, all designs differed significantly from *Hands*, which performed the worst. Overall, *Shape* was rated highest, followed by *Air*, and then *Eyes*.

4 Conclusion

We evaluated four keyboard designs using the Dvorak layout. Designs (a) and (d) were equally efficient, while (d) had better usability and lower Physical Demand. Design (b) increased Physical Demand, Effort, and Frustration. Design (c) was less physically demanding but had greater Mental Demand and the highest error rate. Lower typing speeds may stem from unfamiliarity with Dvorak. Our results underscore the need for keyboard stability to enhance typing speed and accuracy. We suggest using a stable keyboard close to the user with focus-enhancing features, like a bifocal display, to improve VR text input. This study advances understanding of how VR keyboard designs affect text input performance and user satisfaction, aiding the development of more effective VR interfaces.

5 Acknowledgements

The infrastructure for this work is funded by the European Union – NextGenerationEU project *MAXVR-INFRA* and the Flemish government.

References

[1] John Brooke. 1996. *SUS: A "Quick and Dirty" Usability Scale.* CRC Press.

[2] Sandra G. Hart and Lowell E. Staveland. 1988. *Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research.* Human Mental Workload, Vol. 52. North-Holland, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[3] Florian Kern, Florian Niebling, and Marc Erich Latoschik. 2023. Text Input for Non-Stationary XR Workspaces: Investigating Tap and Word-Gesture Keyboards in Virtual and Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 29, 5 (05 2023), 2658–2669. https://doi.org/10.1109/TVCG.2023.3247098

[4] Robin Kinkead. 1975. Typing Speed, Keying Rates, and Optimal Keyboard Layouts. *Proceedings of the Human Factors Society Annual Meeting* 19, 2 (1975), 159–161. https://doi.org/10.1177/154193127501900203 arXiv:https://doi.org/10.1177/154193127501900203

[5] Marius Schenkluhn, Christian Peukert, Anke Greif-Winzrieth, and Christof Weinhardt. 2023. Does One Keyboard Fit All? Comparison and Evaluation of Device-Free Augmented Reality Keyboard Designs. In *29th ACM Symposium on Virtual Reality Software and Technology* (, Christchurch, New Zealand,) *(VRST '23)*. ACM, Christchurch New Zealand, Article 5, 11 pages. https://doi.org/10.1145/3611659.3615692

A Measurement & Statistics Report

Table 1: Measures, statistics and questionnaire results of the performed study. RTLX and SUS ranges from 0–100. Significance markers * for $p < 0.05$, ** for $p < 0.01$ and * for $p < 0.001$.**

Metric/scales	Air Mean (*SD*)	Hand Mean (*SD*)	Eyes Mean (*SD*)	Shape Mean (*SD*)	Repeated Measures ANOVA or Friedman test	Post-hoc tests (p, Z, r)
Entry rate [WPM]	6.51 (1.62)	5.56 (1.47)	6.70 (1.65)	5.33 (1.48)	$F(3, 45) = 9.04$, $p < 0.001$ ***	Air-Eyes ($p < 0.001$ ***, $Z = 4.07$, $r = 0.25$), Air-Hands ($p < 0.001$ ***, $Z = 3.29$, $r = 0.21$), Eyes-Shape ($p < 0.001$ ***, $Z = 4.46$, $r = 0.28$), Shape-Hands ($p < 0.001$ ***, $Z = -3.70$, $r = 0.23$)
Word error rate (WER) [%]	1.24 (4.35)	3.82 (17.58)	1.43 (5.19)	1.64 (5.58)	$\chi^2(3) = 0.45$, $p = 0.931$	
Keystrokes per character [KSPC]	1.11 (0.12)	1.14 (0.15)	1.24 (0.22)	1.09 (0.14)	$\chi^2(3) = 22.82$, $p < 0.001$ ***	Air-Eyes ($p < 0.001$ ***, $Z = -3.71$, $r = 0.23$), Eyes-Shape ($p < 0.001$ ***, $Z = -4.67$, $r = 0.29$), Eyes-Hands ($p = 0.0025$ **, $Z = -3.00$, $r = 0.19$)
Raw NASA-TLX: Mental Demand	29.7 (21.0)	38.8 (27.7)	47.2 (23.8)	35.0 (18.9)	$\chi^2(3) = 7.46$, $p = 0.059$	
Raw NASA-TLX: Physical Demand	56.2 (28.0)	75.6 (20.1)	32.5 (23.0)	37.2 (24.8)	$\chi^2(3) = 31.61$, $p < 0.001$ ***	Air-Eyes ($p = 0.0201$ *, $Z = 2.31$, $r = 0.29$), Air-Shape ($p = 0.0389$ *, $Z = 2.06$, $r = 0.26$), Air-Hands ($p = 0.0288$ *, $Z = -2.18$, $r = 0.27$), Eyes-Hands ($p < 0.001$ ***, $Z = 4.01$, $r = 0.50$), Shape-Hands ($p < 0.001$ ***, $Z = 3.91$, $r = 0.49$)
Raw NASA-TLX: Temporal Demand	25.6 (19.7)	27.8 (19.4)	22.2 (14.0)	29.1 (18.4)	$\chi^2(3) = 2.46$, $p = 0.483$	
Raw NASA-TLX: Performance	33.8 (21.3)	40.9 (19.4)	41.9 (20.6)	28.4 (20.1)	$\chi^2(3) = 6$, $p = 0.112$	
Raw NASA-TLX: Effort	46.6 (24.9)	64.1 (22.1)	52.5 (20.8)	35.9 (19.2)	$\chi^2(3) = 19.01$, $p < 0.001$ ***	Air-Hands ($p = 0.0429$ *, $Z = -2.02$, $r = 0.25$), Eyes-Shape ($p = 0.0287$ *, $Z = -2.18$, $r = 0.27$), Shape-Hands ($p < 0.001$ ***, $Z = 3.47$, $r = 0.43$)
Raw NASA-TLX: Frustration	44.7 (30.8)	59.1 (32.4)	38.1 (23.7)	29.1 (22.3)	$\chi^2(3) = 16.98$, $p < 0.001$ ***	Eyes-Hands ($p = 0.0337$ *, $Z = 2.12$, $r = 0.26$), Shape-Hands ($p = 0.0114$ *, $Z = 2.50$, $r = 0.31$)
System usability scale Score	78.3 (17.6)	56.4 (22.7)	74.8 (14.1)	80.5 (15.1)	$\chi^2(3) = 12.45$, $p = 0.006$ **	Air-Hands ($p = 0.0059$ **, $Z = 2.70$, $r = 0.34$), Eyes-Hands ($p = 0.0154$ *, $Z = -2.40$, $r = 0.30$), Shape-Hands ($p = 0.0020$ **, $Z = -3.00$, $r = 0.38$)
Easiest (1) to hardest (4) to use	2.44 (0.81)	3.75 (0.45)	2.56 (1.09)	1.44 (0.63)	$\chi^2(3) = 26.66$, $p < 0.001$ ***	Air-Shape ($p < 0.001$ ***, $Z = 3.37$, $r = 0.60$), Air-Hands ($p < 0.001$ ***, $Z = -4.02$, $r = 0.71$), Eyes-Shape ($p = 0.0031$ **, $Z = -2.94$, $r = 0.52$), Eyes-Hands ($p < 0.001$ ***, $Z = 3.42$, $r = 0.61$), Shape-Hands ($p < 0.001$ ***, $Z = 4.98$, $r = 0.88$)
Easiest (1) to hardest (4) to understand	1.25 (0.58)	3.06 (0.85)	2.94 (0.93)	2.75 (1.06)	$\chi^2(3) = 20.48$, $p < 0.001$ ***	Air-Eyes ($p < 0.001$ ***, $Z = -4.28$, $r = 0.76$), Air-Shape ($p < 0.001$ ***, $Z = -3.90$, $r = 0.69$), Air-Hands ($p < 0.001$ ***, $Z = -4.56$, $r = 0.81$)
Best (1) to worst (4) for typing	2.19 (0.54)	3.69 (0.48)	2.44 (1.26)	1.69 (0.95)	$\chi^2(3) = 20.85$, $p < 0.001$ ***	Air-Shape ($p = 0.0263$ *, $Z = 2.17$, $r = 0.38$), Air-Hands ($p < 0.001$ ***, $Z = -4.70$, $r = 0.83$), Eyes-Hands ($p = 0.0030$ **, $Z = 2.99$, $r = 0.53$), Shape-Hands ($p < 0.001$ ***, $Z = 4.44$, $r = 0.78$)
Most (1) to least (4) annoying	2.88 (0.89)	1.44 (0.51)	2.25 (1.24)	3.44 (0.63)	$\chi^2(3) = 21.23$, $p < 0.001$ ***	Air-Hands ($p < 0.001$ ***, $Z = 4.03$, $r = 0.71$), Eyes-Shape ($p = 0.0056$ **, $Z = 2.72$, $r = 0.48$), Shape-Hands ($p < 0.001$ ***, $Z = -4.85$, $r = 0.86$)
Most (1) to least (4) pleasant	2.50 (0.82)	3.81 (0.40)	2.06 (1.12)	1.62 (0.62)	$\chi^2(3) = 25.73$, $p < 0.001$ ***	Air-Shape ($p = 0.0044$ **, $Z = 2.96$, $r = 0.52$), Air-Hands ($p < 0.001$ ***, $Z = -4.30$, $r = 0.76$), Eyes-Hands ($p < 0.001$ ***, $Z = 4.15$, $r = 0.73$), Shape-Hands ($p < 0.001$ ***, $Z = 5.01$, $r = 0.89$)

B Sentences

Table 2: The practice sentences are the four sentences that participants had to type with the mechanical physical keyboard for practice. The study sentences are the five sentences that participants had to type in VR.

Practice Sentences	Study Sentences
rent is paid at the beginning of the month	a big scratch on the tabletop
ask not what your country can do for you	the quick brown fox jumped
east west north south	my car always breaks in the winter
up down left right	dolphins leap high out of the water
	video camera with a zoom lens