The Effect of Target Depth on Performance of Multi-directional Tapping Task in Virtual Reality

Haejun Kim
KAIST
Daejeon, South Korea
vedonefor@kaist.ac.kr

Yuhwa Hong
KAIST
Daejeon, South Korea
yuhwahong@kaist.ac.kr

Jihae Yu
Kangwon National University
Chuncheon, South Korea
vievanitas@kangwon.ac.kr

Shuping Xiong*
KAIST
Daejeon, South Korea
shupingx@kaist.ac.kr

Woojoo Kim*
Kangwon National University
Chuncheon, South Korea
woojoo.kim@kangwon.ac.kr

ABSTRACT

While widely used to evaluate 2D pointing performance, adapting the multi-directional tapping task (ISO/TS 9241-411) to virtual reality (VR) poses challenges, particularly in addressing target depth. This study examines how depth affects user performance in the multi-directional tapping task in VR. We conducted a within-subject experiment with 20 participants, investigating the effect of various depths (0.5–100 m for Raycasting; 0.3–0.6 m for Virtual Hand) under consistent visual angles. Results showed that Raycasting performance remained stable beyond 2 m but degraded significantly at 0.5 m, while Virtual Hand performed best between 0.4 and 0.5 m and declined at closer and farther depths. These findings suggest that target depth strongly influences selection performance even when visual angles remain consistent, underscoring the need for considering standardized depth parameters in VR pointing protocols. We also provide evidence-based recommendations for implementing depth parameters in future VR studies using the multi-directional tapping task.

CCS CONCEPTS

• **Human-centered computing → HCI theory, concepts and models**; **HCI design and evaluation methods**; **Pointing**.

KEYWORDS

Virtual reality, Multi-directional tapping task, Target selection, Target depth, Fitts' law

ACM Reference Format:
Haejun Kim, Yuhwa Hong, Jihae Yu, Shuping Xiong, and Woojoo Kim. 2025. The Effect of Target Depth on Performance of Multi-directional Tapping Task in Virtual Reality. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26-May 1, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3706599.3719893

*Corresponding authors.

CHI EA '25, April 26-May 1, 2025, Yokohama, Japan

ACM ISBN 979-8-4007-1395-8/2025/04
https://doi.org/10.1145/3706599.3719893

1 INTRODUCTION

In the field of human-computer interaction (HCI), researchers have relied on standardized evaluation methods to assess interaction techniques and input devices. The multi-directional tapping task, formalized in ISO/TS 9241-411 [14] (previously known as ISO 9241-9 [13]) and also known as 2D Fitts' law task [37], has emerged as a particularly valuable tool for measuring user performance in traditional graphical interfaces [10, 36, 44]. As virtual reality (VR) technology matures and finds applications across entertainment, training, healthcare, and other domains [39, 42, 48], researchers have increasingly adapted this standardized task to evaluate interaction in 3D virtual environments [45]. A standard task for 3D interfaces has yet to be established, which is why the multi-directional tapping task is widely used in VR. However, this transition from 2D to VR interfaces has introduced significant challenges in maintaining consistent experimental protocols, particularly regarding the configuration of target depth parameters.

The core challenge lies in systematically integrating depth—a dimension absent in traditional 2D interfaces—into VR tapping tasks. Since the multi-directional tapping task was designed to evaluate 2D pointing movements, the corresponding standard provides guidelines for target placement in 2D spaces but offers no specific guidance for addressing depth in virtual environments. Consequently, researchers have made widely varying choices in target positioning, ranging from within arm's reach to far-field distances [7]. These inconsistencies in depth configuration make it difficult to compare results across studies, as performance variations might stem from methodological differences rather than meaningful patterns in user behavior. The lack of standardized depth parameters particularly affects the evaluation of fundamental VR interaction techniques, such as Raycasting and Virtual Hand [9, 30].

Further complicating matters is the relationship between target depth and size in VR environments. Unlike 2D interfaces, where the target width remains visually constant, VR can place objects in a vast, potentially infinite space far from the user, amplifying perspective effects so that targets appear smaller at greater depths. Researchers have addressed this challenge differently: some maintain consistent target width regardless of depth, while others keep depth fixed and control target visual angles, defined as the angular size of the target as perceived by the user based on its distance and physical size (see Table 1). These divergent approaches to target size configuration, combined with varying depth parameters, create a complex web of methodological differences that undermines the comparability of research findings. Without standardized protocols for handling these parameters, we lack a reliable foundation for building cumulative knowledge about VR interaction performance.

Our research addresses these standardization challenges by conducting a systematic investigation of the effect of target depth while ensuring consistent target visual angles in the multi-directional tapping task within VR. Through experiments with both Raycasting and Virtual Hand techniques, we examine how different configurations of these parameters affect key performance metrics such as movement time, error rate, and throughput. We focus on establishing evidence-based guidelines for implementing these parameters in future VR studies. Our goal is to contribute to the development of more standardized protocols that enable meaningful comparisons across different research efforts while maintaining the methodological rigor established by ISO/TS 9241-411.

2 RELATED WORK

2.1 Fitts' Law and the Multi-directional Tapping Task

Fitts' law has served as a cornerstone in HCI research since its introduction in 1954. While initially developed for 1D tasks, researchers have successfully extended its application to 2D interfaces, leading to significant advances in input device evaluation and interface design [26, 41]. This mathematical model has proven particularly valuable in standardizing performance measurements across different input devices and interaction techniques, with its integration into the ISO/TS 9241-411 standard marking a crucial milestone in HCI research methodology. The standard multi-directional tapping task has emerged as a widely adopted tool for evaluating pointing devices, offering researchers a systematic approach to measuring user performance [6, 28, 31].

This Fitts' law-based task has been continuously improved through important changes in how it is conducted, particularly in the calculation and interpretation of performance metrics. Researchers have introduced sophisticated approaches to measuring effective target width (W_e) and effective target amplitude (A_e), which account for the actual distribution of movement endpoints rather than nominal target parameters [24, 41]. These refinements have enhanced the task's ability to capture subtle differences in user performance, leading to more reliable and comparable results across studies. The standardized nature of the task has facilitated comprehensive evaluations of various input devices, from traditional mice and touchpads to emerging technologies like eye-tracking and gesture-based interfaces [16, 18, 43, 47].

Recent adaptations of the multi-directional tapping task have expanded beyond traditional interface evaluations. With the emergence of VR technologies, researchers have begun adapting these established evaluation methods for 3D spaces, leading to new challenges in implementing and interpreting this standardized task.

2.2 Fitts' Law-Based Studies in Virtual Reality

While numerous studies have attempted to extend Fitts' law to 3D selection tasks in VR environments [11, 23], these investigations have largely overlooked the independent effect of depth on user performance. Most studies have incorporated depth as a compound variable affecting both target width and amplitude, rather than isolating its unique impact on selection behavior.

Research exploring Fitts' law in VR has generally followed three distinct approaches (see Table 1). The first approach maintains consistent depth while varying visual angles, similar to traditional 2D studies. The second approach fixes depth while varying target width, which limits our understanding of how depth influences selection performance. The third approach varies both target width and depth, potentially confounding the effects of depth and apparent size.

While these studies have advanced our understanding of selection behavior in VR, they share a critical limitation: the inability to isolate depth's effect from other variables. Real-world VR applications require users to interact with objects at various depths and sizes, yet existing models typically do not separate depth with apparent target width, making it difficult to predict performance across different scenarios. Our study addresses this limitation by systematically investigating depth effects while maintaining consistent visual angles, allowing us to better understand depth's independent influence on selection behavior.

3 EXPERIMENT

3.1 Participants and Experimental Setup

A total of 20 participants (8 females and 12 males) with a mean age of 22.45 years (SD = 3.21) participated in the study. All participants were right-handed individuals who had normal or corrected-to-normal vision, with arm lengths including hand measuring less than 0.6 m. Most participants had prior experience with VR systems. The research protocol received approval from the University Institutional Review Board (IRB NO.: KWNUIRB-2024-09-006-001). Before participating in the study, all participants provided written informed consent.

The experimental setup is illustrated in Figure 1. All experiments were conducted using a Meta Quest Pro VR headset. Participants were seated in a chair and used the controller in their right hand to complete the experimental task, with instructions to maintain a stable posture. The experimental environment was developed using Unity (version 2022.3.51f1).

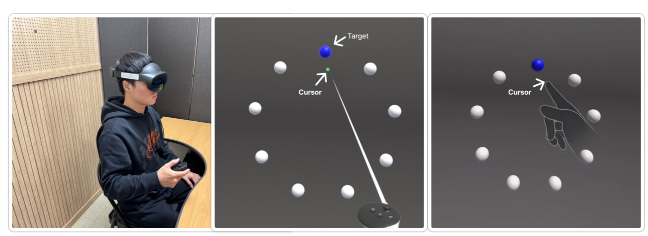

Figure 1: The experimental environment (left) and the multi-directional tapping task setup with Raycasting (middle) and Virtual Hand (right) in the virtual environment. The center point of the target ring was equal to the eye height [33]. For a demo video, please refer to: https://youtu.be/uY9oBxAVD3A.

Parameters	Methods for Adjusting Target Size	References
Target Visual Angle	Adjusting target visual angle without varying target depth	[21, 40, 46]
Target Width	Adjusting target width without varying target depth	[4, 5, 12, 29]
Target Width & Target Depth	Adjusting both target width and target depth	[2, 3, 5, 15, 34, 35, 38]

Table 1: Fitts' law-based VR studies and their methods for adjusting target size regarding target width, visual angle, and depth control.

3.2 Experimental Design

This study employed a 4 × 3 × 3 within-subjects design for each of the two selection techniques: Raycasting and Virtual Hand. Raycasting is a selection technique where users point at targets using a virtual ray projected from a controller or hand, allowing for distant object interaction. Virtual Hand, on the other hand, involves directly manipulating virtual objects within arm's reach by simulating the user's hand movements in the virtual environment. The following factors and levels were investigated (Figure 2):

- **Target Depth**: 4 levels were tested for each interaction technique. For Virtual Hand, the depths were 0.3, 0.4, 0.5, and 0.6 m. For Raycasting, the depths were 0.5, 2, 10, and 100 m. These depth conditions were chosen to cover a wide range of binocular disparities [19], from high at close depths to minimal at far depths.
- **Target Visual Angle**: 3 levels were tested: 1°, 2.5°, and 4°.
- **Target Amplitude**: 3 levels were tested: 25°, 35°, and 45°.

The combination of these independent variables resulted in a wide range of index of difficulty (ID) values ranging from 2 to 6. This design led to a total of 4 depths × 3 visual angles × 3 amplitudes = 36 unique conditions for each interaction technique. Each participant completed 9 trials per condition, resulting in a total of 36 conditions × 9 trials × 2 interaction techniques = 648 data points per participant.

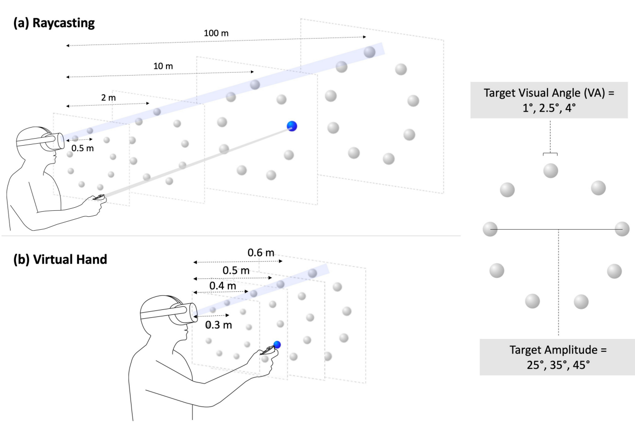

Figure 2: Experimental design for two selection techniques: (a) Raycasting and (b) Virtual Hand.

3.3 Experimental Procedure

Participants were randomly assigned to one of two groups in a counterbalanced design. One group performed Virtual Hand followed by Raycasting, while the other group completed the tasks in reverse order. For each technique, participants completed two sessions of the multi-directional tapping task under 36 test conditions presented in random order. The first session was considered a practice run and was excluded from subsequent data analysis. The multi-directional tapping task involved participants tapping alternately between multiple targets arranged in a circular pattern, requiring precise and rapid pointing movements [14]. After completing all tasks, a follow-up interview was conducted to gather participants' perceptions of how depth affected their performance, as well as their reasoning behind these perceptions. The experiment took around 40–50 minutes to complete, and each participant received a monetary reward of KRW 30,000 (~ USD 20) for completing the experiment.

3.4 Data Analysis

The analysis utilized three primary performance metrics: movement time (MT), error rate (ER), and effective throughput (TP) [25]. Consistent with prior methodologies in pointing and crossing studies [17, 27], all trial data, including those with errors, were included to ensure a comprehensive analysis with the derivation of W_e.

To identify outliers, the interquartile range (IQR) method was employed, defining potential outliers as data points exceeding 1.5 times the IQR. As a result, data from two participants with exceptionally high error rates (ER > 0.4) were excluded. Following data cleaning, linear regression analyses were performed for each depth condition to explore the relationship between MT and ID. Unadjusted R^2 values were used to assess the goodness of fit for the regression models.

Statistical analyses were conducted using repeated measures ANOVA (RM-ANOVA) with three within-subject factors: target depth, visual angle, and amplitude. When Mauchly's test indicated a violation of sphericity, Greenhouse-Geisser corrections were applied. All analyses were performed at $\alpha = .05$ using Python 3.10.11 and R 4.4.1.

4 RESULTS

While our RM-ANOVA examined three factors, we focus our analysis specifically on the depth to address our core research questions. A full RM-ANOVA table of all performance measures is provided in the Appendix A.

4.1 Performance

The results revealed distinct performance patterns across different depths for both interaction techniques (see Figure 3 and Table 2). For Raycasting, the shortest depth (0.5 m) resulted in a significantly longer MT (M = 1.074 s, SD = 0.370) compared to other depths, whereas no significant differences in MT were observed among the 2, 10, and 100 m depths. This was also evident in the ER, with the 0.5 m condition showing a significantly higher ER (M = 0.143, SD = 0.159) compared to other depths. Similarly, TP was generally consistent across most depths, except for the 0.5 m condition which demonstrated notably lower performance (M = 3.693 bit/s, SD = 1.036).

For Virtual Hand, moderate performance variations were observed across depths, with 0.4 and 0.5 m conditions demonstrating particularly favorable results. The 0.6 m condition had the longest MT (M = 0.809 s, SD = 0.212), followed by the 0.5 m condition (M = 0.770 s, SD = 0.203). In contrast, the 0.3 m (M = 0.728 s, SD = 0.200) and 0.4 m (M = 0.729 s, SD = 0.193) conditions had comparably shorter MT. ER was highest at 0.3 m (M = 0.148, SD = 0.168) and lowest at 0.5 m (M = 0.105, SD = 0.136). For TP, peak performance was observed at 0.5 m (M = 4.976 bit/s, SD = 1.309), closely followed by the 0.4 m condition (M = 4.848 bit/s, SD = 1.230).

Additionally, we found significant two-way interaction effects between target depth and target visual angle on ER (Figure 4). Error patterns varied substantially for small visual angles (1°) compared to larger ones. In particular, the combination of small visual angles and short depths led to notably higher error rates. No other significant two-way or three-way interaction effects related to target depth were identified among the factors.

4.2 Model Fitting

To investigate depth's influence on Fitts' law model fitness, we calculated the R^2 values across four linear regression models for each interaction method (Figure 5). For Raycasting, strong linear relationships between MT and ID were found in all depth conditions. The shortest depth (0.5 m) showed the lowest R^2 value (R^2 = 0.97), while both 10 and 100 m conditions achieved slightly better fits (R^2 = 0.99). All depth conditions maintained exceptional model fits with R^2 > 0.97, demonstrating strong adherence to Fitts' law predictions regardless of depth. In contrast, Virtual Hand exhibited more noticeable variations across depths compared to Raycasting. The closest condition (0.3 m) demonstrated the strongest fit (R^2 = 0.96), while other depths showed lower R^2 values: 0.6 m (R^2 = 0.87), 0.4 m (R^2 = 0.84), and 0.5 m (R^2 = 0.82). This suggests that Fitts' law models' predictive accuracy is more depth-dependent in Virtual Hand than in Raycasting.

4.3 Subjective Feedback

In the Raycasting condition, a significant majority (80%) of participants reported perceiving noticeable performance differences across the varying depth conditions. Participants described a range of experiences at different depths. For instance, P14 stated: *'The farthest distance was the easiest. I think it's because it required less hand movement,'* which underscores a key benefit of interacting with distant targets in this condition. Conversely, when targets were positioned closer, participants reported adapting their physical behavior. P6 mentioned: *'When targets were too close, I tended to lean my body backward.'* This observation, along with reports from other participants indicating increased hand movements and postural adjustments at closer distances, suggests consistent strategies to physically compensate for proximity. These subjective experiences align with our objective measurements, which revealed performance degradation at the closest distance (0.5 m).

In the Virtual Hand condition, 85% of participants reported perceiving performance differences across the varied depth conditions. Notably, participants consistently identified challenges at both extremes of the depth range. P15 highlighted the physical strain associated with reaching for distant targets, stating: *'When targets were farther away, my arm would hurt and I felt less accurate.'* P11 offered a similar observation: *'As the distance increased, I had to move my head more and stretch my arm further, which was tiring.'* On the other hand, interacting with targets at the closest distance presented distinct challenges. P9 observed: *'When targets were too close, my hand would accidentally trigger errors while moving between targets.'* These subjective reports correlate with our objective findings, which demonstrated optimal performance at intermediate distances (0.4-0.5 m) and performance degradation at both the closest and farthest distances.

5 DISCUSSION

5.1 Raycasting

In the Raycasting condition, our results revealed that depth primarily affects performance at close ranges (0.5 m), while maintaining consistent performance from medium (2 m) to far distances (100 m). This pattern can be attributed to several key factors. First, although targets were kept consistent in visual angle across conditions, at the closest depth (0.5 m), targets were positioned near the participant's head. This made Raycasting uncomfortable at lower hand positions, forcing participants to raise their hands, which increased muscle fatigue and resulted in longer MT and lower TP [20]. Second, with respect to ER, the proximity between the controller and target at 0.5 m introduced two challenges: small rotations of the controller caused exaggerated cursor movements, making precise control more difficult, and the raised hand position led to greater instability and tremors [22, 32]. Finally, the consistent performance observed at greater depths (2, 10, and 100 m) suggests that once targets are beyond a certain range, depth has minimal impact on Raycasting performance. This stability can be attributed to the nature of Raycasting [8], where the physical movement required remains relatively constant regardless of target depth.

5.2 Virtual Hand

Virtual Hand exhibited more pronounced depth-dependent variations in performance. Our results indicate an optimal performance range between 0.4-0.5 m, with significant degradation observed at both closer (0.3 m) and farther (0.6 m) distances. MT showed a proportional relationship with depth, as shorter depths naturally required less physical movement [1]. This relationship reflects the direct mapping of physical to virtual hand movements inherent to Virtual Hand interaction. While TP did not show significant variations across depths, ER revealed an intriguing trend: despite consistent target visual angles, errors decreased from 0.3 m to 0.5 m as the absolute target width increased with depth. However, errors increased again at 0.6 m, likely due to physical discomfort from extended reaching movements [1], as supported by Penumudi et al., who found that target locations in a multi-directional tapping task

Techniques	Movement Time			Error Rate			Throughput		
	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
Raycasting	8.100	***<0.001	0.025	4.634	**0.013	0.014	23.656	***<0.001	0.077
Virtual Hand	18.803	***<0.001	0.044	4.356	**0.014	0.017	2.599	0.079	0.011

Table 2: F, p, and η_p^2 values from the RM-ANOVA results showing the effects of depth on performance metrics. $^{*}p< 0.05$, $^{**}p< 0.01$, $^{***}p< 0.001$.

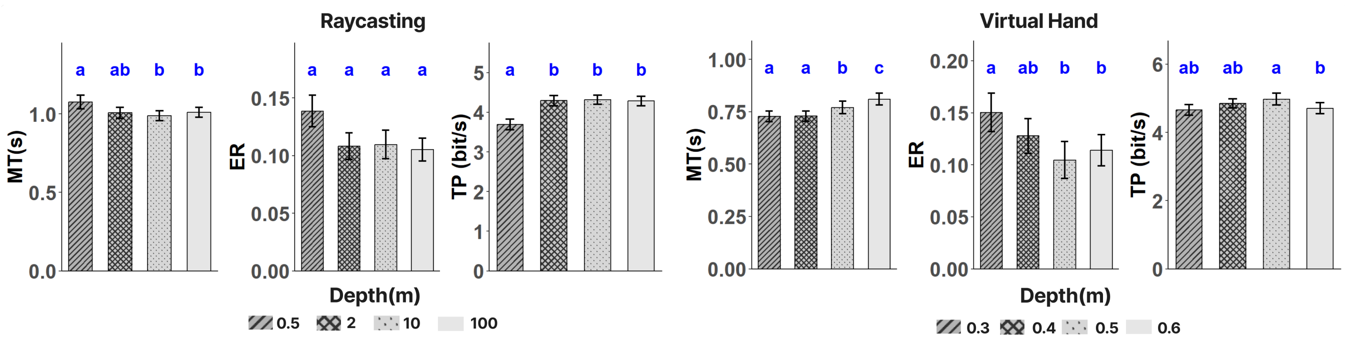

Figure 3: Mean performance measures by target depth for each selection technique. Error bars indicate standard error of the mean. The letters represent Bonferroni post-hoc groupings. Bars that do not share the same letter are significantly different.

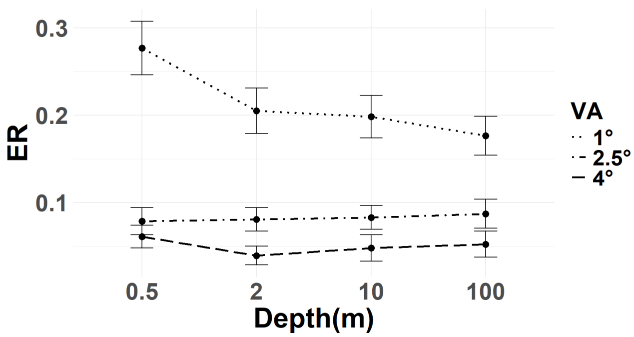

Figure 4: Mean error rate by target depth and visual angle. Error bars indicate standard error of the mean.

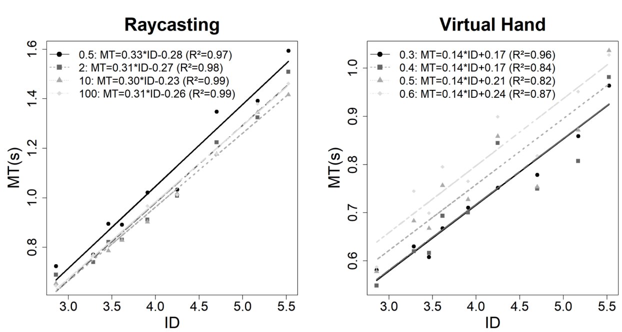

Figure 5: Linear regression plots between MT and ID, and R^2 by target depth for each selection technique.

in VR significantly impacted biomechanical strain on the neck and shoulders [33]. This U-shaped error pattern suggests an optimal interaction zone that balances target accessibility with arm comfort [22].

5.3 Implications

These findings have important implications for implementing multi-directional tapping tasks in VR. For evaluations of pointing with Raycasting, our results suggest that testing with target depths beyond 2 m, which aligns with the original purpose of Raycasting to select distant objects, may not be critical when using a multi-directional tapping task, as performance remains consistent across these conditions. However, if testing at close depths (0.5 m) is necessary, researchers should consider the significant performance degradation observed at this range, as it highlights the need to account for depth conditions in the task design and interpretation of results.

For evaluations of pointing with Virtual Hand, target depths should be set around 0.4-0.5 m when evaluating baseline performance, as this represents the optimal interaction zone. It is crucial to note that performance metrics are highly sensitive to depth changes, requiring careful depth standardization across studies. These findings suggest that standardized protocols for the multi-directional tapping task in VR should incorporate depth-specific considerations, even when maintaining consistent visual angles, to ensure reliable and comparable evaluation results across studies.

5.4 Limitations and Future Work

It's important to acknowledge several limitations that may influence the interpretation of our findings. Firstly, the effect sizes (η_p^2)

observed for depth-related effects were comparatively modest, particularly when contrasted with the effects of visual angle and amplitude. This is likely attributable to the multi-directional tapping task itself, where, as predicted by Fitts' law, target size and inter-target distance are primary determinants of performance variability.

Secondly, the scope of our conclusions is inherently tied to the specific VR hardware and virtual environment design employed in this study. The technical specifications of the Meta Quest Pro, including its field of view, resolution, focal plane, and stereo rendering, may have influenced our observations regarding depth perception. While the headset's specifications and technological implementation represent a contemporary dominant approach, results obtained using different VR headsets may diverge. Furthermore, our intentionally minimalist virtual environment, designed to isolate the impact of depth, lacks the rich depth cues found in real-world settings, such as shadows, textures, perspective, and occlusion. This may limit the direct applicability of our findings to typical VR applications where these cues interact to shape user perception and performance. While this controlled environment provides a valuable baseline for understanding the fundamental role of depth in selection behavior, more complex, application-specific environments may yield different performance patterns.

Finally, our study examined a limited set of discrete depth levels, which may not fully represent the nuanced effects of depth perception within VR. Future research could enhance our understanding by exploring selection behavior in visually richer environments that incorporate multiple depth cues, investigating target selection across a wider range of depth variations, and accounting for individual differences in depth perception. These extensions would more accurately reflect real-world VR applications and contribute to the development of more comprehensive models of selection behavior in virtual environments. Additionally, cross-device validation with a variety of VR headsets would bolster the generalizability of our findings across the broader VR ecosystem."

6 CONCLUSION

In this study, we systematically investigated how target depth affects user performance in VR multi-directional tapping tasks while maintaining consistent target visual angles, aiming to develop evidence-based guidelines that foster more standardized protocols in future VR research. The findings of this study highlight the significant role of target depth in influencing the performance of the multi-directional tapping task, even when the target visual angle remains constant. For Raycasting, depth increases beyond 2 m showed no notable impact on performance, but performance degraded at shorter depths. Conversely, the Virtual Hand exhibited sensitivity to even minor depth variations as small as 0.1 m, with the best performance observed at approximately 0.5 m. These results underscore the importance of carefully considering target depth when designing or performing multi-directional tapping tasks in VR environments. Furthermore, the majority of participants were able to discern performance differences between varied depth conditions, despite the visual angle of the targets being identical. This emphasizes the perceptual and interactive nuances introduced by depth for using the multi-directional tapping task in VR.

ACKNOWLEDGMENTS

This work was supported by Innovative Human Resource Development for Local Intellectualization program through the Institute of Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korean government(MSIT)(IITP-2024-RS-2023-00260267) and the National Research Foundation of Korea(NRF) grants funded by the Korean government(MSIT)(RS-2024-00343882, RS-2023-00242528).

REFERENCES

[1] Anil Ufuk Batmaz, Mayra Donaji Barrera Machuca, Duc Minh Pham, and Wolfgang Stuerzlinger. 2019. Do Head-Mounted Display Stereo Deficiencies Affect 3D Pointing Tasks in AR and VR?. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 585–592. https://doi.org/10.1109/VR.2019.8797975

[2] Anil Ufuk Batmaz, Aunnoy K Mutasim, and Wolfgang Stuerzlinger. 2020. Precision vs. Power Grip: A Comparison of Pen Grip Styles for Selection in Virtual Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 23–28. https://doi.org/10.1109/VRW50115.2020.00012

[3] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2020. Effect of Fixed and Infinite Ray Length on Distal 3D Pointing in Virtual Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3334480.3382796

[4] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2022. Effective Throughput Analysis of Different Task Execution Strategies for Mid-Air Fitts' Tasks in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* 28, 11 (2022), 3939–3947. https://doi.org/10.1109/TVCG.2022.3203105

[5] Anil Ufuk Batmaz, Kangyou Yu, Hai-Ning Liang, and Wolfgang Stuerzlinger. 2022. Improving Effective Throughput Performance using Auditory Feedback in Virtual Reality. In *Proceedings of the 2022 ACM Symposium on Spatial User Interaction* (Online, CA, USA) *(SUI '22)*. Association for Computing Machinery, New York, NY, USA, Article 18, 11 pages. https://doi.org/10.1145/3565970.3567702

[6] T R Beelders and P J Blignaut. 2012. Using eye gaze and speech to simulate a pointing device. *Eye Tracking Research and Applications Symposium (ETRA)* (2012), 349–352. https://doi.org/10.1145/2168556.2168634

[7] Joanna Bergström, Tor-Salve Dalsgaard, Jason Alexander, and Kasper Hornbæk. 2021. How to Evaluate Object Selection and Manipulation in VR? Guidelines from 20 Years of Studies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 533, 20 pages. https://doi.org/10.1145/3411764.3445193

[8] Doug A. Bowman, Donald B. Johnson, and Larry F. Hodges. 1999. Testbed evaluation of virtual environment interaction techniques. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (London, United Kingdom) *(VRST '99)*. Association for Computing Machinery, New York, NY, USA, 26–33. https://doi.org/10.1145/323663.323667

[9] Doug A. Bowman, Ernst Kruijff, Joseph J. LaViola, and Ivan Poupyrev. 2004. *3D User Interfaces: Theory and Practice.* Addison Wesley Longman Publishing Co., Inc., USA.

[10] Kyeong Kyun Cheong, Insoo Kim, Sung Kwon Park, and Yong Jin Park. 2011. User performance measures for evaluating interactive TV pointing devices. *IEEE Transactions on Consumer Electronics* 57 (12 2011), 1236–1244. Issue 3. https://doi.org/10.1109/TCE.2011.6018879

[11] Logan D. Clark, Aakash B. Bhagat, and Sara L. Riggs. 2020. Extending Fitts' law in three-dimensional virtual environments with current low-cost virtual reality technology. *International Journal of Human-Computer Studies* 139 (2020), 102413. https://doi.org/10.1016/j.ijhcs.2020.102413

[12] Kristen Grinyer and Robert Teather. 2023. Low-Fi VR Controller: Improved Mobile Virtual Reality Interaction via Camera-Based Tracking. In *Proceedings of the 2023 ACM Symposium on Spatial User Interaction* (Sydney, NSW, Australia) *(SUI '23)*. Association for Computing Machinery, New York, NY, USA, Article 5, 12 pages. https://doi.org/10.1145/3607822.3614517

[13] ISO. 2000. ISO 9241-9. Ergonomic requirements for office work with visual display terminals (VDTs) — Part 9: Requirements for non-keyboard input devices.

[14] ISO. 2012. ISO/TS 9241-411. International standard: ergonomics of human-system interaction — Part 411: Evaluation methods for the design of physical input devices.

[15] Jin-Woo Jeong and Jae-Yeop Jeong. 2024. Effect of Onset Position of Ray Casting in Virtual Reality. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 116, 7 pages. https://doi.org/10.1145/3613905.3650905

[16] Marcelo Archajo José and Roseli De Deus Lopes. 2015. Human-computer interface controlled by the lip. *IEEE Journal of Biomedical and Health Informatics* 19 (1

2015), 302–308. Issue 1. https://doi.org/10.1109/JBHI.2014.2305103

[17] Nobuhito Kasahara, Yosuke Oba, Shota Yamanaka, Wolfgang Stuerzlinger, and Homei Miyashita. 2023. Throughput and Effective Parameters in Crossing. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 282, 9 pages. https://doi.org/10.1145/3544549.3585817

[18] Keiko Katsuragawa, Krzysztof Pietroszek, James R Wallace, and Edward Lank. 2016. Watchpoint: Freehand pointing with a smartwatch in a ubiquitous display environment. *Proceedings of the Workshop on Advanced Visual Interfaces AVI* 07-10-June-2016 (12 2016), 128–135. https://doi.org/10.1145/2909132.2909263

[19] Jonathan W. Kelly. 2023. Distance Perception in Virtual Reality: A Meta-Analysis of the Effect of Head-Mounted Display Characteristics. *IEEE Transactions on Visualization and Computer Graphics* 29, 12 (2023), 4978–4989. https://doi.org/10.1109/TVCG.2022.3196606

[20] Eunjee Kim and Gwanseob Shin. 2018. Head rotation and muscle activity when conducting document editing tasks with a head-mounted display. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 62. SAGE Publications Sage CA: Los Angeles, CA, 952–955.

[21] Woojoo Kim and Shuping Xiong. 2022. ViewfinderVR: configurable viewfinder for selection of distant objects in VR. *Virtual Real.* 26, 4 (Dec. 2022), 1573–1592. https://doi.org/10.1007/s10055-022-00649-z

[22] Xiaolong Lou, Xiangdong Li, Preben Hansen, and Zhipeng Feng. 2020. An Empirical Evaluation on Arm Fatigue in Free Hand Interaction and Guidelines for Designing Natural User Interfaces in VR. In *Virtual, Augmented and Mixed Reality. Design and Interaction*, Jessie Y. C. Chen and Gino Fragomeni (Eds.). Springer International Publishing, Cham, 313–324.

[23] Xiaolong Lou, Xuanbai Song, Xiaofan Hu, Mengzhen Ma, Lili Fu, Xinran Zhuang, and Xueer Fan. 2024. An Extended Fitts' Law Model for Hand Interaction Evaluation in Virtual Reality: Accounting for Spatial Variables in 3D Space and the Arm Fatigue Influence. *International Journal of Human–Computer Interaction* 0, 0 (2024), 1–27. https://doi.org/10.1080/10447318.2024.2443708 arXiv:https://doi.org/10.1080/10447318.2024.2443708

[24] I Scott MacKenzie. 1989. A Note on the Information-Theoretic Basis for Fitts' Law. *Journal of Motor Behavior* 21 (12 1989), 323–330. Issue 3. https://doi.org/10.1080/00222895.1989.10735486

[25] I. Scott Mackenzie. 2015. Fitts' Throughput and the Remarkable case of touch-based target Selection. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9170 (2015), 238–249. https://doi.org/10.1007/978-3-319-20916-6_23/FIGURES/11

[26] I Scott MacKenzie and William Buxton. 1992. Extending Fitts' law to two-dimensional tasks. *Conference on Human Factors in Computing Systems - Proceedings* (1992), 219–226. https://doi.org/10.1145/142750.142794

[27] I. Scott MacKenzie and Poika Isokoski. 2008. Fitts' throughput and the speed-accuracy tradeoff. *Conference on Human Factors in Computing Systems - Proceedings* (2008), 1633–1636. https://doi.org/10.1145/1357054.1357308

[28] Shuhei Matsuyama and Mitsuhiko Karashima. 2017. A study on characteristics of hand gesture pointing operation versus mouse pointing operation: A comparison of velocity waves of operation time between mouse pointing and hand gesture pointing with two kinds of control-display ratio. *Communications in Computer and Information Science* 713 (2017), 170–176. https://doi.org/10.1007/978-3-319-58750-9_24/FIGURES/4

[29] Ken McAnally and Guy Wallis. 2023. Effects of Auditory Feedback on Visually-Guided Movement in Real and Virtual Space. *International Journal of Human-Computer Interaction* 40 (05 2023), 1–10. https://doi.org/10.1080/10447318.2023.2216091

[30] Mark R. Mine. 1995. *Virtual Environment Interaction Techniques.* Technical Report. USA.

[31] Daniel Natapov and I Scott MacKenzie. 2010. The trackball controller: Improving the analog stick. *Future Play 2010: Research, Play, Share - International Academic Conference on the Future of Game Design and Technology* (2010), 175–182. https://doi.org/10.1145/1920778.1920803

[32] Dan R. Olsen and Travis Nielsen. 2001. Laser pointer interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, USA) *(CHI '01)*. Association for Computing Machinery, New York, NY, USA, 17–22. https://doi.org/10.1145/365024.365030

[33] Sai Akhil Penumudi, Veera Aneesh Kuppam, Jeong Ho Kim, and Jaejin Hwang. 2020. The effects of target location on musculoskeletal load, task performance, and subjective discomfort during virtual reality interactions. *Applied Ergonomics* 84 (2020), 103010. https://doi.org/10.1016/j.apergo.2019.103010

[34] Max Pfeiffer and Wolfgang Stuerzlinger. 2015. 3D Virtual Hand Selection with EMS and Vibration Feedback. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI EA '15)*. Association for Computing Machinery, New York, NY, USA, 1361–1366. https://doi.org/10.1145/2702613.2732763

[35] Duc-Minh Pham and Wolfgang Stuerzlinger. 2019. Is the Pen Mightier than the Controller? A Comparison of Input Devices for Selection in Virtual and Augmented Reality. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 35, 11 pages. https://doi.org/10.1145/3359996.3364264

[36] Alexandros Pino, Evangelos Tzemis, Nikolaos Ioannou, and Georgios Kouroupetroglou. 2013. Using Kinect for 2D and 3D Pointing Tasks: Performance Evaluation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 8007 LNCS (2013), 358–367. Issue PART 4. https://doi.org/10.1007/978-3-642-39330-3_38

[37] Paul M Pitts. 1954. THE INFORMATION CAPACITY OF THE HUMAN MOTOR SYSTEM IN CONTROLLING THE AMPLITUDE OF MOVEMENT. *Journal of Experimental Psychology* 47 (1954). Issue 6.

[38] Adrian Ramcharitar and Robert J. Teather. 2018. EZCursorVR: 2D Selection with Virtual Reality Head-Mounted Displays. In *Proceedings of the 44th Graphics Interface Conference* (Toronto, Canada) *(GI '18)*. Canadian Human-Computer Communications Society, Waterloo, CAN, 123–130. https://doi.org/10.20380/GI2018.17

[39] Mickael Sereno, Xiyao Wang, Lonni Besançon, Michael J. McGuffin, and Tobias Isenberg. 2022. Collaborative Work in Augmented Reality: A Survey. *IEEE Transactions on Visualization and Computer Graphics* 28, 6 (June 2022), 2530–2549. https://doi.org/10.1109/TVCG.2020.3032761

[40] Rongkai Shi, Yushi Wei, Yue Li, Lingyun Yu, and Hai-Ning Liang. 2023. Expanding Targets in Virtual Reality Environments: A Fitts' Law Study . In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE Computer Society, Los Alamitos, CA, USA, 615–618. https://doi.org/10.1109/ISMAR-Adjunct60411.2023.00132

[41] R William Soukoreff and I Scott MacKenzie. 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *International Journal of Human-Computer Studies* 61 (12 2004), 751–789. Issue 6. https://doi.org/10.1016/J.IJHCS.2004.09.001

[42] Hidayat Ullah, Selvakumar Manickam, Muath Obaidat, Shams Ul Arfeen Laghari, and Mueen Uddin. 2023. Exploring the Potential of Metaverse Technology in Healthcare: Applications, Challenges, and Future Directions. *IEEE Access* 11 (2023), 69686–69707. https://doi.org/10.1109/ACCESS.2023.3286696

[43] Eduardo Velloso, Jason Alexander, Andreas Bulling, and Hans Gellersen. 2015. Interactions under the desk: A characterisation of foot movements for input in a seated position. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9296 (2015), 384–401. https://doi.org/10.1007/978-3-319-22701-6_29/TABLES/3

[44] Jacob O Wobbrock, Kristen Shinohara, and Alex Jansen. 2011. The effects of task dimensionality, endpoint deviation, throughput calculation, and experiment design on pointing measures and models. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA). Association for Computing Machinery, 1639–1648. https://doi.org/10.1145/1978942.1979181

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

[46] Difeng Yu, Hai-Ning Liang, Xueshi Lu, Kaixuan Fan, and Barrett Ens. 2019. Modeling endpoint distribution of pointing selection tasks in virtual reality environments. *ACM Trans. Graph.* 38, 6, Article 218 (Nov. 2019), 13 pages. https://doi.org/10.1145/3355089.3356544

[47] Xuan Zhang and I Scott MacKenzie. 2007. Evaluating Eye Tracking with ISO 9241 - Part 9. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 4552 LNCS (2007), 779–788. Issue PART 3. https://doi.org/10.1007/978-3-540-73110-8_85

[48] M. Zyda. 2005. From visual simulation to virtual reality to games. *Computer* 38, 9 (2005), 25–32. https://doi.org/10.1109/MC.2005.297

A A FULL RM-ANOVA TABLE

Technique	Source	Movement Time F	Movement Time p	Movement Time η_p^2	Error Rate F	Error Rate p	Error Rate η_p^2	Throughput F	Throughput p	Throughput η_p^2
Raycasting	**D**	8.100	***<0.001	0.025	4.634	**0.013	0.014	23.656	***<0.001	0.077
	VA	232.724	***<0.001	0.605	47.780	***<0.001	0.282	102.425	***<0.001	0.208
	A	216.531	***<0.001	0.149	3.861	*0.038	0.014	1.376	0.265	0.003
	D:A	1.527	0.214	0.005	0.603	0.645	0.006	0.898	0.471	0.005
	D:VA	1.813	0.139	0.005	2.805	*0.029	0.024	2.280	0.060	0.011
	VA:A	3.366	*0.026	0.007	0.979	0.410	0.006	0.549	0.644	0.003
	D:A:VA	1.074	0.376	0.006	1.231	0.300	0.018	1.236	0.289	0.015
Virtual Hand	**D**	18.803	***<0.001	0.044	4.356	*0.014	0.017	2.599	0.079	0.011
	VA	125.428	***<0.001	0.259	43.918	***<0.001	0.209	32.477	***<0.001	0.086
	A	106.873	***<0.001	0.195	27.056	***<0.001	0.042	7.337	**0.004	0.021
	D:A	2.387	0.051	0.009	1.260	0.292	0.006	0.954	0.447	0.005
	D:VA	0.544	0.676	0.003	1.350	0.265	0.007	1.888	0.117	0.013
	VA:A	3.080	*0.029	0.008	4.730	*0.013	0.015	2.193	0.098	0.012
	D:A:VA	0.944	0.472	0.007	0.506	0.779	0.005	1.346	0.239	0.015

Table 3: A RM-ANOVA table of all performance variables. D = Target Depth, VA = Target Visual Angle, A = Target Amplitude.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$