Comparing Controller-Free Pointing Techniques Across Depth for 2D Selection in Augmented Reality

Samiha Sultana
Carleton University
Ottawa, Canada
SamihaSultana@cmail.carleton.ca

J. Felipe Gonzalez
Carleton University
Ottawa, Canada
johannavila@cunet.carleton.ca

Robert J. Teather
Monash University
Melbourne, Australia
rob.teather@monash.edu

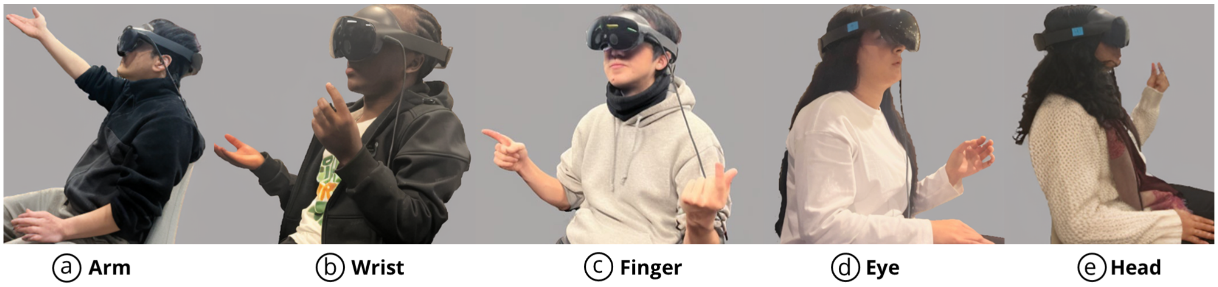

Figure 1: The five controller-free pointing techniques for study: (a) Arm, (b) Wrist, (c) Finger, (d) Eye, and (e) Head. Each was tested at 2, 6, and 10 meters of depth to evaluate movement time, error rate, throughput, and workload in AR 2D selection.

ABSTRACT

This paper presents a systematic evaluation of five controller-free pointing techniques for 2D target selection in AR, using ISO 9241-411. We compared them across multiple depths (2 m, 6 m, 10 m) in terms of movement time, accuracy, throughput, and workload (NASA TLX). Head- and eye-based pointing significantly outperformed the hand-based methods (Finger, Wrist, and Arm); Head input was the most accurate and remained the most consistent across depth. Depth significantly impacted performance, with complex interactions with target size and distance. Our results offer a comprehensive empirical basis for selecting appropriate controller-free techniques in depth-varying AR tasks.

CCS CONCEPTS

- **Human-centered computing** → **Interaction techniques**; *User studies.*

KEYWORDS

Pointing, Augmented Reality, Virtual Reality, Fitts' law

ACM Reference Format:
Samiha Sultana, J. Felipe Gonzalez, and Robert J. Teather. 2026. Comparing Controller-Free Pointing Techniques Across Depth for 2D Selection in Augmented Reality. In *GI 2026: Graphics Interface, June 9–12, 2026, Waterloo, Canada.* ACM, New York, NY, USA, 11 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Remote pointing facilitates interaction in both augmented reality (AR) and virtual reality (VR), allowing users to aim at distant points or select and manipulate objects beyond physical reach [3, 50, 77]. Traditionally, remote pointing requires controllers to manipulate the origin and direction of a virtual ray; pressing a button afterward completes the interactions, such as confirming object selection or locomotion [10, 36]. However, recent advances in hand- [14, 62], head-, and eye-tracking [38, 59] support new **controller-free** pointing techniques including arm-, wrist-, finger-, head-, and eye-based methods. These present competitive alternatives, potentially eliminating the need for physical controllers [11].

Prior studies have explored the benefits and limitations of these techniques. Hand-based input (e.g., arm, wrist, and finger) is intuitive and supports proprioception [7, 41], but it can yield fatigue [18, 27]. Wrist-based pointing is less fatiguing [14], but also less precise, especially for very far targets [53, 55]. The finger is widely used for manipulation, gestures, and UI interaction [9, 34, 46], but due to tracking limitations it is rarely used for remote pointing and is underexplored [57]. Head-based pointing offers stable, hands-free interaction [40], while eye-gaze is fast [38, 59], but it is sensitive to calibration errors and involuntary eye movements [8].

Unfortunately, prior work differs widely in focus and methodology. There are evaluations of specific combinations of pointing technique and selection indication mechanisms, such as gaze with dwell, pinch, or button click [8, 52]. Other work evaluated only limited subsets of modalities, such as gaze, foot, and head [49], or compared controllers with gaze or finger input [20, 34, 48]. Several studies explore compound interactions (e.g., gaze with hand input), but do not isolate the individual contributions of each modality [43, 67, 76]. Most prior research evaluates an input modality *relative to* a hand-held controller [16, 20, 58, 69], rather than directly comparing controller-free modalities against each other. To date

---

GI '26, June 9–12, 2026, Waterloo, Canada
© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *GI 2026: Graphics Interface, June 9–12, 2026, Waterloo, Canada*, https://doi.org/XXXXXXX.XXXXXXX.

there are no systematic evaluations isolating hand-, finger-, or wrist-based pointing performance [82]. Comparing these input modalities (i.e., finger, wrist, arm, eye and head) systematically is important because each one offers distinct strengths and limitations (e.g., movement time, error rate, and workload) that are often confounded in techniques that combine multiple input modalities.

Another difficulty in comparing previous results lies in the variety of study tasks, such as menu selection in AR [8] or in virtual and real-world environments [85]. ISO 9241-411 [1] is based on Fitts' law [21], and improves study comparability through use of a standardized task and measures. Fitts' law is a robust predictive model of pointing performance and is widely applied in VR [2, 30, 60, 63] and, to a lesser extent, in AR [12]. According to Fitts' law, pointing performance is influenced by target distance and width in both VR [15, 35, 49, 52, 76] and AR [6, 38, 85]. Most Fitts' law evaluations in 3D environments overlook the influence of target depth in 3D selection tasks [2]. Yet, in 3D environments target depth influences pointing performance due to both perspective scaling and since the ray's angular precision degrades with distance [29, 36, 37]. Stereo displays, including all commercially available head-mounted displays (e.g., Meta Quest Pro), are also subject to vergence-accommodation conflicts, impacting depth perception and compromising interaction [6, 78].

We present a study addressing these gaps via a rigorous, systematic, and generalizable evaluation. Most prior pointing research focused on VR [22, 30, 33, 60, 63, 84], with comparatively fewer studies in AR [12, 32, 54]. Although early work reported differences between VR and AR [32, 54], more recent findings suggest that this gap may be smaller with improved hardware [6]. We therefore conduct an up-to-date, systematic comparison of controller-free pointing techniques in AR. Our research question is: ***What are the performance and user experience differences between controller-free remote pointing techniques in AR?***

We employed the ISO 9241-411 [1] standard selection task to compare accuracy and completion time of five controller-free remote pointing techniques: Finger, Wrist, Arm, Eye, and Head (Figure 1). Additionally, we assessed user workload using the NASA-TLX questionnaire [26]. By comparing these interaction modalities, we offer insights into which techniques are best suited to AR tasks with varying spatial demands.

In summary, the main contributions of our work include:

- A systematic evaluation of controller-free pointing techniques (Finger, Wrist, Arm, Eye, and Head) in AR.
- A thorough examination of the impact of target depth on selection performance with these techniques.
- An analysis of pointing efficiency and the speed/accuracy trade-off across input modalities using Fitts' law metrics.

2 BACKGROUND

2.1 Pointing in AR/VR

Remote pointing supports reaching specific points or objects beyond arm's reach [31, 36, 65] in both VR and AR [3, 39]. It supports distal interaction while reducing the need for locomotion [81]. It involves visually identifying a target, **pointing** with a cursor (e.g., via ray-casting), and **confirming** selection through actions like a button press, a pinch gesture, or dwell [8, 10, 52].

Prior work suggests that using separate hands for pointing and confirmation improves performance: Sindhupathiraja et al. [70] reported better results when users pointed with the dominant hand and confirmed with a non-dominant-hand pinch than when the same hand was used for both actions. This division helps mitigate the Heisenberg effect, where executing the confirmation action can degrade pointing accuracy [56, 80]. Consequently, all confirmations in our study used a non-dominant-hand pinch gesture.

While remote pointing historically relied on controllers, modern devices (e.g., Meta Quest 3 [47]) enable *controller-free* pointing through head-, hand-, and eye-tracking, supporting modalities such as Head, Eye, Arm, Wrist, and Finger [14, 38, 59, 62]. Prior work examined multimodal combinations (e.g., eye+head or eye+hand) and confirmation mechanisms (pinch, click, gestures, dwell) [41, 43, 52, 59, 68, 77]. Other studies focused on aiming dynamics (e.g., gaze behavior or target properties) [15, 48, 67]. However, these studies usually emphasize combinations or dynamics rather than fundamental comparisons of standalone pointing techniques.

Several studies have directly compared gaze and head pointing against other inputs. Minakata *et al.* [49] compared gaze, head, and foot pointing and reported head outperforming gaze. Lin *et al.* [41] found hand better for tracking, while gaze was more effective for discrete pointing. Hansen *et al.* [24] compared mouse, head, and gaze, and Qian and Teather [61] reported head-based pointing outperforming eye and the combination of eye and head.

In contrast, Arm-, Wrist-, and Finger-based pointing remains underexplored as standalone selection techniques. Although common for navigation or gesture interaction in VR [30, 33, 63, 84], they have rarely been evaluated in isolation. Early work explored finger/wrist/arm using wearables [57], but not in VR and is less representative of current AR tracking. Studies using handheld controllers [16, 20, 69] differ fundamentally from controller-free modalities. A survey of 106 papers on VR/AR selection and manipulation [82] reported no direct comparisons of these hand-based modalities as standalone selection techniques.

Moreover, most pointing research targets VR [22, 30, 33, 60, 63, 84], with fewer studies in AR [12, 32, 54]. While early work reported differences between VR and AR [32, 54], more recent findings suggest smaller gaps, potentially due to improved hardware [6]. This motivates an up-to-date, systematic comparison of controller-free pointing techniques in AR.

2.2 Selection and Fitts' Law

Target selection techniques, including remote pointing, are commonly evaluated using ISO 9241-411 [1], and Fitts' law [24, 44, 51, 57]. Fitts' law predicts movement time (MT) to acquire a target as:

$$MT = a + b \log_2 \left(\frac{A}{W} + 1 \right) \tag{1}$$

where A is the movement amplitude (i.e., distance to the target), W the target width (*i.e.* , target size). The log term is the index of difficulty, or ID, which increases with higher amplitudes and/or smaller targets. The coefficients a and b are empirically derived through linear regression between recorded average movement time and ID. To mitigate speed-accuracy trade-off biases [83], an accuracy adjustment [17] is applied to calculate the effective Index of Difficulty, ID_e. Using ID_e, throughput (TP) is calculated as:

$$TP = \frac{ID_e}{MT}, \quad \text{where} \quad ID_e = \log_2\left(\frac{A_e}{W_e} + 1\right) \tag{2}$$

Here, effective width, $W_e = 4.133 \cdot SD_x$ where SD_x is the standard deviation of distance between selection coordinates and the target centre. This accounts for variability in selection coordinates. W_e is an adjusted target size where 96% of selections would have hit the target (i.e., 4% error rate). This accuracy adjustment facilitates comparing throughput across studies with varying error rates [83]. Effective amplitude, A_e, is the average distance from the cursor start position to the end position in each trial. Both effective measures better capture the task participants perform, rather than what is presented. We used Fitts' law to analyze pointing effectiveness to improve comparability to previous research.

2.3 Depth Perception

Perceptual problems persist in AR/VR displays, impacting target acquisition. Notably, the convergence/accommodation mismatch impacts user ability to localize targets in 3D [20, 28]. Target depth — the distance of the interaction plane from the user [31, 64, 72] — also influences pointing performance [29, 36, 37]. Due to perspective, target depth directly relates to target width in display space (i.e., farther targets appear smaller), which in turn affects selection time [29, 35, 37]. Teather and Stuerzlinger [74] found that stereo cursors can help with near-screen targets but degrade performance for deeper ones due to diplopia (i.e., double vision). Recent work [5] in VR using modern headsets like the Quest 3 confirmed that selection performance significantly degrades (i.e., increasing movement time and reducing throughput) as targets move away from the focal plane due to the vergence-accommodation conflict. Triantafyllidis and Li [75] highlight existing research gaps in VR/AR pointing, which notably includes that the impact of depth perception issues is underexplored. We therefore evaluate performance across multiple target depths for several controller-free pointing modalities.

3 POINTING TECHNIQUE DESIGN SPACE

Pointing in AR is influenced by choice of pointing technique, target depth, and, per Fitts' law, target size and distance. We discuss the contributions of each below.

3.1 Pointing Techniques

We selected five controller-free pointing techniques to cover the main tracking capabilities of modern AR HMDs. This included three hand-based ray origins (Finger, Wrist, Arm) alongside two hands-free ray origins (Head, Eye). This facilitates controlled evaluation of how different ray origins affect selection performance under identical conditions (e.g., confirmation and cursor).

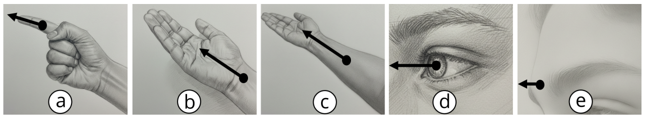

Figure 2: The five pointing techniques, depicting ray origin and direction: (a) Finger, (b) Wrist, (c) Arm, (d) Eye, (e) Head.

1. Finger: By aligning the index finger with the target (Figure 2a), participants rely on very fine motor control in the finger [57]. The selection ray is cast from the fingertip, allowing rapid pointing.

2. Wrist: The pointer ray originates from the user's dominant wrist (Figure 2b), leveraging fine motor control at the wrist joint [13, 14]. The device's built-in SDK[1] handles tracking, and wrist orientation controls the ray's direction. Wrist was included because the ray direction is directly steered by wrist rotation, allowing fine-grained control through the wrist joint itself.

3. Arm: The pointer ray originates from the distal forearm (Figure 2c) and the ray direction is controlled by the user's arm orientation. This relies on gross motor control with the ray anchored at the lower arm region [14]. We included both Arm and Wrist because they rely on different control mechanisms. In Arm, the ray is primarily steered by forearm/arm orientation, and wrist rotation does not affect cursor movement.

4. Eye: Targets were selected via gaze using the device eye tracker[1], see Figure 2d. The ray direction was determined by the eye gaze vector provided by the device SDK[1], which represents the user's gaze in world space based on the tracked eye pose.

5. Head: The pointer ray originates from the midpoint between the user's eyes (Figure 2e), following the approach in Movement SDK[2]. This frees the hands [38] and is useful when arms are fatigued or engaged in other tasks [53].

No filtering (e.g., smoothing) was applied to any of the five techniques. This ensured that our evaluation captures the baseline performance of each method without the confounding influence of a filtering algorithm [45] or the latency they can introduce [73].

3.2 Target Width and Amplitude

According to Fitts' law, target size and distance influence selection task difficulty (ID) and thus target acquisition time. It is thus customary to vary target width (W) and movement amplitude (A) to present a wide range of IDs, in accordance with ISO 9241-411 [1, 71]. This provides the advantage of comparing different pointing techniques/devices across a range of task difficulties (e.g., easy selection tasks vs. hard selection tasks), enhancing experimental external validity. Previous work mostly considered linear target width and movement amplitude, i.e., measured in meters or pixels [21, 25, 42, 55, 66, 72]. Therefore, we used target width of 0.2 m, 0.3 m, and 0.4 m and for target amplitude we used 1.5 m, 2.5 m, and 3.5 m. The combination of these width and amplitude values results in 9 distinct ID levels, ranging from 2.24 to 4.21 bits (see eqn. 1). Gori et al. [23] cautioned that confounding effects in Fitts' law studies can arise when different amplitude/width combinations yield the same ID, especially in designs with geometrically distributed amplitudes and widths. In our study, amplitude and width values were not geometrically distributed, and the nine amplitude and width combinations yielded nine distinct ID values.

3.3 Target Depth

In our study, target depth is the perpendicular distance from the user to the interaction plane where the targets reside [31, 64, 72]. There

---

[1] https://developer.oculus.com/documentation/unity/unity-isdk-interaction-sdk-overview/

[2] https://developer.oculus.com/documentation/unity/move-overview/

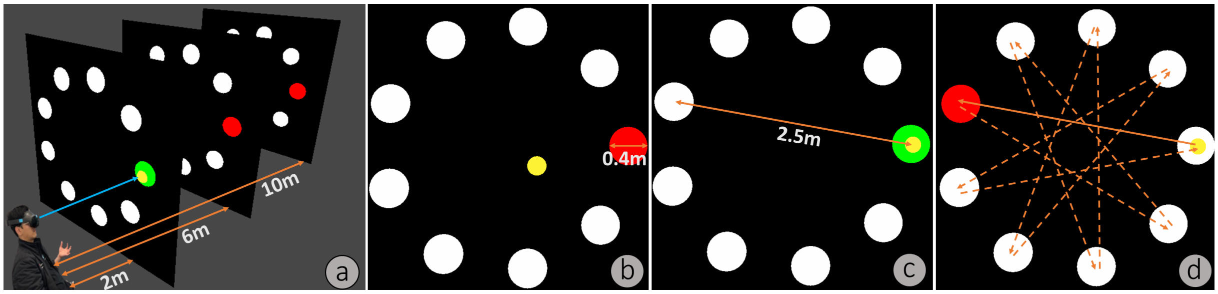

Figure 3: (a) Study setup with different target depth conditions, (b) yellow cursor approaching target with 0.4m target width, (c) yellow cursor colliding with target with green visual cue and the target amplitude is 2.5m, and (d) After confirming the selection the next target turns red. The solid arrow shows the path to the current target, while dashed arrows represent the sequential order in which the user must select the 9 targets.

are known issues relating to depth perception [28, 75], especially across different input methods [75]. We therefore investigate 3 depth values (2 m, 6 m, and 10 m away from the viewer) across the five pointing techniques. We maintained constant target widths in meters across depths (similar to previous Fitts' law studies [4, 15, 70, 74]) rather than fixed angular widths. While this design decision inherently varies visual angle, it prioritizes ecological validity: in real-world AR scenarios, physical objects do not dynamically scale with user distance. Table 1 summarizes the corresponding visual angles for each amplitude/width/depth combination.

		Depth (m)								
		2			**6**			**10**		
Amplitude (m)		1.5	2.5	3.5	1.5	2.5	3.5	1.5	2.5	3.5
Width (m)	**0.2**	5.36	4.86	4.31	1.89	1.87	1.83	1.14	1.14	1.13
	0.3	8.03	7.28	6.46	2.84	2.80	2.75	1.71	1.71	1.69
	0.4	10.70	9.70	8.61	3.79	3.74	3.67	2.29	2.27	2.26

Table 1: Angular widths in degrees for each combination of width (W), amplitude (A), and depth (D).

4 METHODOLOGY

We conducted a systematic evaluation of five controller-free pointing techniques for AR target selection, evaluating both target depth and selection task difficulty.

4.1 Participants

We recruited 20 participants (11 men, 9 women), aged 19 to 48 years (mean = 27.95, *SD* = 8.51), from the local university. All participants were right-handed, had (corrected-to-) normal vision, with no known motor impairments. Of the 20 participants, 8 had no prior experience with AR/VR. The study was conducted in Canada; each participant provided consent in accordance with Carleton University's Research Ethics Board protocol, and received $15 CAD for their participation.

4.2 Apparatus

We used a Meta Quest Pro headset for its built-in hand- and eye-tracking capabilities, high-resolution (1800 x 1920 pixels per eye) displays, and wide field of view (106° horizontal, 96° vertical). Participants sat in a quiet laboratory setting, free from distractions, ensuring consistent experimental conditions.

We implemented an ISO 9241-411 reciprocal target selection task in Unity3D. Circular 2D targets were arranged in a ring at three depths (2 m, 6 m, 10 m) with widths (0.2 m, 0.3 m, 0.4 m) and amplitudes (1.5 m, 2.5 m, 3.5 m), yielding nine *ID* levels (2.24–4.21 bits). The software presented a 20 cm yellow cursor (as seen in Figure 3 and 4), which resided in the plane of the targets where the selection ray intersected that plane. For each trial, the current target changed red while the other targets remained white. When the cursor intersected the current target, it turned green to indicate that it was ready for selection. Selection was confirmed with a pinch gesture by the non-dominant-hand thumb and forefinger. If the cursor was outside the target, the system registered confirmations as a miss; however, the trial continued until the participant successfully selected the target, after which the next target in the sequence would become active and turn red. The target sequence is seen in Figure 3d. We used Meta's OVRHand SDK to obtain tracked poses for the pointing techniques (see section 3.1) and logged all trial data via Firebase Realtime Database[3].

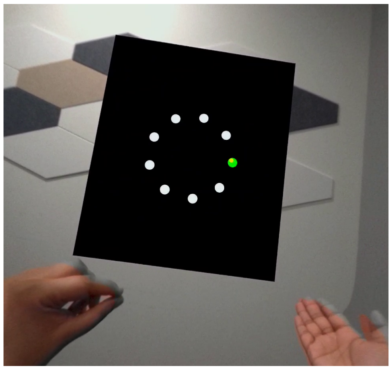

Figure 4: Augmented Reality View of the task from participant's perspective for Wrist. When the yellow cursor intersects the red target (seen in figure 3), the target turns green to provide visual feedback that it is ready for selection.

4.3 Procedure

Upon arrival, participants provided written informed consent prior to participating. They were then introduced to the Quest Pro

---

[3] https://firebase.google.com/products/realtime-database

headset and the general study objectives. Before starting the actual trials for each pointing technique, participants completed at least 10 practice trials to familiarize themselves with the pointing technique and the confirmation gesture. For each condition, we instructed participants to primarily use the intended pointing technique to control the ray via the specified joint or pose. If they were observed using a joint other than the target joint during the Arm, Wrist, and Finger conditions, we gently reminded them to only use the specified joint. Eye-tracking calibration was performed prior to the eye pointing condition. Head motion was not constrained, as users naturally rotate their heads during target acquisition [55], particularly when target amplitude is high [38].

In each trial, a target appeared at a specified depth, width, and amplitude combination. Participants were instructed to select the target as quickly and accurately as possible using the currently assigned pointing technique. After acquiring the target (i.e., aligning the pointer with the target), they performed a pinch gesture with their non-dominant hand to confirm selection.

Participants completed a NASA TLX questionnaire and provided comments after each technique. They were allowed to rest and briefly remove the headset before starting the next technique if needed. Upon completing all trials with each selection technique, participants completed a short demographic survey and ranked the techniques. We then thanked them for their time and compensated them at the end of the experiment.

4.4 Design

Our study followed a 5 × 3 × 3 × 3 within-subjects design with the following independent variables and levels:

- *Pointing Technique:* Finger, Wrist, Arm, Eye, Head
- *Width:* 0.2 m, 0.3 m, 0.4 m
- *Amplitude:* 1.5 m, 2.5 m, 3.5 m
- *Depth:* 2 m, 6 m, 10 m

For each combination of independent variables, the user had to perform 9 trials. Each trial involved selecting a single target. Pointing technique order was counterbalanced using a Latin square design to mitigate order effects. The 9 combinations of Amplitude and Width yielded 9 distinct *ID*s, representing varying task difficulty. For each Pointing Technique, the order of *ID* and Depth were randomized. Each participant completed 5 (techniques) × 9 (IDs) × 3 (depths) × 9 (trials per ID) = 1215 trials, for a total of 24,300 trials across all 20 participants.

Our dependent variables included:

- **Movement Time:** The duration between selecting the previous target and selecting the current target, in seconds.
- **Error rate:** The percentage of trials where the selection ended outside the target boundary (i.e., missed the target).
- **NASA TLX:** Subjective workload ratings via the NASA TLX questionnaire. We calculated the overall workload using raw NASA-TLX values.
- **Throughput:** We also calculate the **Throughput (TP)** derived from movement time and effective *ID* (via effective width and amplitude) using equation 2.

5 RESULTS

We removed outliers where the movement time exceeded three standard deviations from the mean (1.65% of data), leaving 23,898 trials. We analyzed time and throughput via repeated measures ANOVA with Bonferroni corrections. Sphericity violations were corrected using Greenhouse-Geisser ($\epsilon_{GG} < .75$) or Huynh-Feldt ($\epsilon_{GG} \geq .75$). Error rate and NASA-TLX data were analyzed using the Aligned Rank Transform (ART) ANOVA [19, 79].

In Figures 6, 7, and 8, pairwise significant differences are depicted as horizontal lines where $p < .05$. Squares enclosing the endpoints of these lines indicate conditions that were significantly different from all other conditions. Significant differences across depths are omitted for clarity.

5.1 Movement Time

We found that Fitts’ law accurately models movement time in terms of *ID*, with $R^2 \geq 0.88$ for all techniques, as indicated by the Fitts’ Law regression lines in Figure 5. The Head- and Eye-based pointing techniques offered the fastest movement times and the smallest sensitivity to *ID*, resulting in faster and more stable performance compared to Arm, Wrist, and Finger techniques. There were significant main effects for both pointing technique ($F_{4,76} = 57.73$, $p < .001$, $\eta^2 = 0.75$) and depth ($F_{2,38} = 63.19$, $p < .001$, $\eta^2 = 0.77$) on movement time.

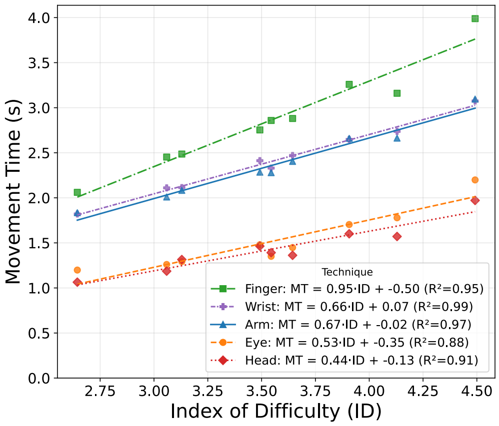

Figure 5: Movement Time vs Index of Difficulty

Pairwise comparisons of the pointing techniques revealed that both Head (mean = 1.43 s, *SD* = 0.07 s) and Eye (mean = 1.55 s, *SD* = 0.10 s) were significantly faster than Arm (mean = 2.38 s, *SD* = 0.10 s), Wrist (mean = 2.43 s, *SD* = 0.15 s), and Finger (mean = 2.93 s, *SD* = 0.14 s), with all comparisons significant at $p < .001$. For *Depth*, all pairwise comparisons between all three levels—2 m (mean = 1.91 s, *SD* = 0.08 s), 6 m (mean = 2.03 s, *SD* = 0.08 s), and 10 m (mean = 2.50 s, *SD* = 0.12 s) were significantly different.

The interaction effect Pointing Technique × Depth was also significant ($F_{4.06,152} = 6.97$, $p < .001$, $\eta^2 = 0.27$). Pairwise comparisons revealed depth-related differences between most pointing techniques, with a few exceptions: for the Arm, there was no significant difference between 2 m (mean = 2.15 s, *SD* =

0.13 s) and 6 m (mean = 2.25 s, *SD* = 0.10 s); for the Head, there was no significant difference between 2 m (mean = 1.45 s, *SD* = 0.07 s) and 6 m (mean = 1.34 s, *SD* = 0.06 s), nor between 2 m and 10 m (mean = 1.52 s, *SD* = 0.10 s). For the Wrist, there was no significant difference between 2 m (mean = 2.28 s, *SD* = 0.14 s) and 6 m (mean = 2.26 s, *SD* = 0.15 s). As seen in Figure 6, depth effects generally followed the main effect trends, with a few notable deviations. At 2 m, Eye was the fastest pointing technique (mean = 1.14 s, *SD* = 0.07 s), significantly outperforming all others, including the Head (mean = 1.45 s, *SD* = 0.07 s). At 2 m depth, the Arm and Finger pointing techniques were not significantly different; however, they were significantly different at the 6 m and 10 m depths. At 10 m depth, the Head pointing technique (mean = 1.50 s, *SD* = 0.10 s) was fastest, significantly outperforming Eye (mean = 2.00 s, *SD* = 0.16 s). Notably, performance with Head was stable regardless of depth; MT with all other pointing techniques increased with greater depth.

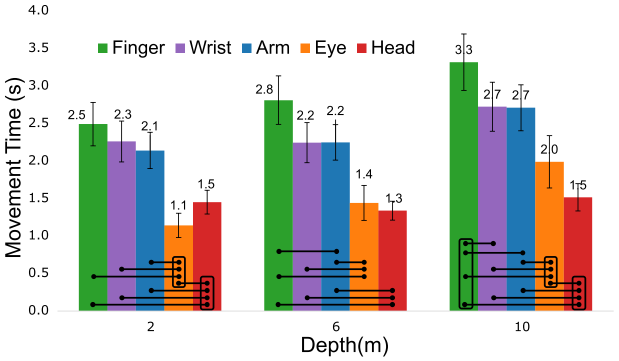

Figure 6: Mean Movement Time per depth per technique. Error bars indicate standard error.

It is also worth noting that amplitude and width influenced the finger and wrist pointing techniques, yielding significant Pointing Technique × Amplitude ($F_{3.96,75.29} = 3.68$, $p < .01$, $\eta^2 = 0.16$) and Pointing Technique × Width ($F_{8,152} = 11.39$, $p < .001$, $\eta^2 = 0.38$) interaction effects. Specifically, at an amplitude of 1.5 m, the performance of Finger and Wrist was not significantly different, while at 2.5 m and 3.5 m, Finger was significantly slower than Wrist. For width, there was no significant difference at the largest width (0.4 m), but Finger was significantly slower than Wrist at the smaller widths of 0.2 m and 0.3 m.

5.2 Error Rates

There were significant main effects on error rate for Pointing Technique ($F_{4,2565} = 170.92$, $p < .001$) and Depth ($F_{4,2565} = 180.21$, $p < .001$). Head had a lower error rate (mean = 5.81 %, *SD* = 0.70 %) than the other four pointing techniques.

Pairwise comparisons revealed significant differences between all technique pairs, except for Arm and Wrist. Eye (mean = 14.27 %, *SD* = 1.53 %) also had a lower error rate than Wrist (mean = 16.91 %, *SD* = 1.39 %), Arm (mean = 17.55 %, *SD* = 1.61 %), and Finger (mean = 24.58 %, *SD* = 1.40 %). Similarly, all depths were significantly different from one another. A significant interaction effect between Pointing Technique and Depth ($F_{8,2565} = 9.94$, $p < .001$), indicated

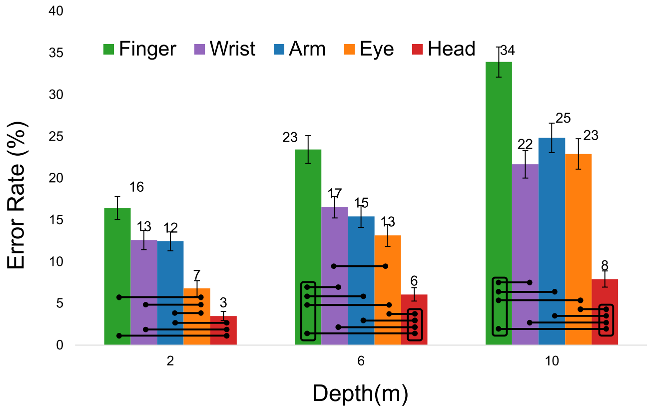

Figure 7: Error Rate per Technique and Depth. Error bars indicate standard error.

Arm error rates were stable between 2 m and 6 m, while Eye and Finger degraded significantly at every depth interval. Head and Wrist remained stable between 6 m and 10 m, suggesting superior stability with greater depth.

5.3 Throughput

There were significant main effects for Pointing Technique ($F_{1.83,34.81} = 56.45$, $\eta^2 = 0.75$) and Depth ($F_{2,38} = 125.58$, $\eta^2 = 0.87$) on throughput. Pairwise comparisons between the Pointing Techniques (see Figure 8) reveal that Head (mean = 2.41 bits/s, *SD* = 0.11 bits/s) and Eye (mean = 2.53 bits/s, *SD* = 0.16 bits/s) offered significantly higher (all $p < .001$) throughput than Arm (mean = 1.48 bits/s, *SD* = 0.08 bits/s), Wrist (mean = 1.47 bits/s, *SD* = 0.09 bits/s) and Finger (mean = 1.21 bits/s, *SD* = 0.06 bits/s). Additionally, it was found that all other pointing techniques had significantly higher throughput than Finger (all $p < .005$).

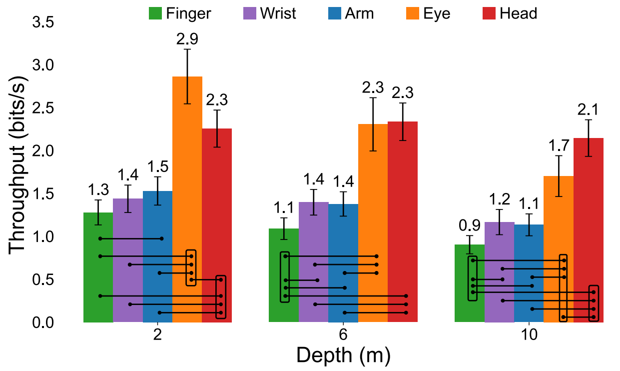

Figure 8: Throughput per technique per depth. Error bars indicate standard error.

For the effect of Depth, pairwise comparisons revealed that pointing from 2 m (mean = 2.00 bits/s, *SD* = 0.09 bits/s) had significantly higher (all $p < .001$) *TP* than both 6 m (mean = 1.89 bits/s, *SD* = 0.08 bits/s) and 10 m (mean = 1.57 bits/s, *SD* = 0.08 bits/s).

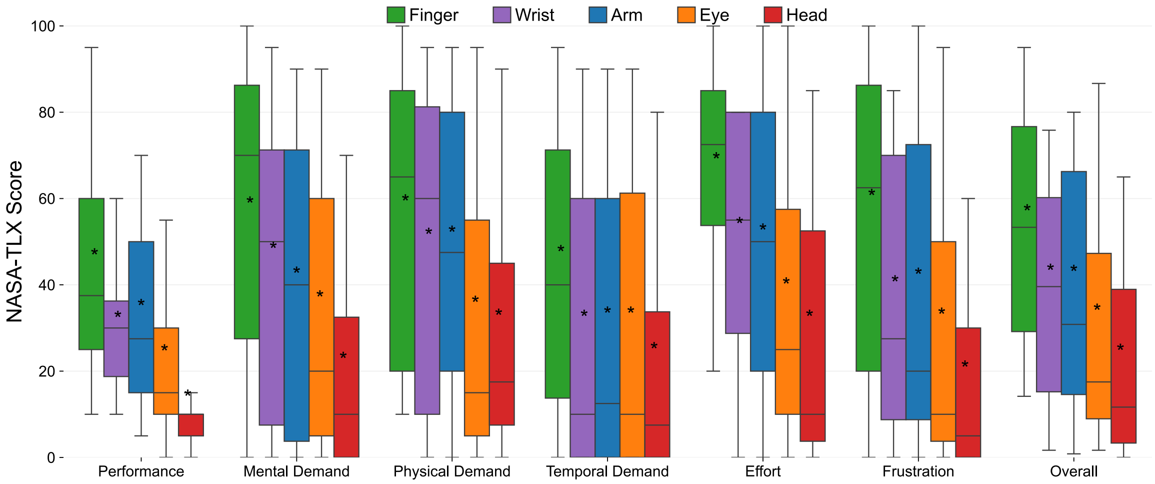

Figure 9: NASA TLX scores vs. criteria per Technique.

The significant interaction effect between Pointing Technique and Depth presented as for the 2 m and 6 m depths, both Eye and Head had higher throughput than Arm, Wrist, and Finger, while at 10 m, Head had a higher throughput than Eye, Arm, Wrist, and Finger.

5.4 NASA TLX

We found significant differences between Pointing Techniques across all NASA TLX metrics (Figure 9): Performance ($F_{4,76} = 18.72$, $p < .001$), Mental Demand ($F_{4,76} = 14.6$, $p < .001$), Physical Demand ($F_{4,76} = 6.59$, $p < .001$), Temporal Demand ($F_{4,76} = 6.43$, $p < .001$), Effort ($F_{4,76} = 11.5$, $p < .001$), Frustration ($F_{4,76} = 11.97$, $p < .001$), and Overall Task Load ($F_{4,76} = 16.83$, $p < .001$).

Post-hoc comparisons revealed that Head was consistently rated as the best technique. It outperformed all other techniques in terms of perceived Performance, and showed significantly lower scores in Mental Demand (vs. all), Effort (vs. Arm, Wrist, and Finger), and Frustration (vs. Arm, Wrist, and Finger). For Physical Demand, Head was significantly lower than Finger, better than Arm, but not significantly different from Eye or Wrist.

The Eye-based pointing technique also performed well, with higher perceived Performance than Finger, and lower Mental and Physical Demand than Finger. In contrast, the Finger-based pointing technique was consistently rated as the most demanding technique. It received the highest (i.e., worst) scores in Mental and Temporal Demand, Effort, Frustration, and Overall Task Load, being significantly worse than most or all other techniques in each case.

5.5 Pointing Technique Ranking

Participants ranked the five techniques according to their subjective preference, as illustrated in Figure 10. Head was the most preferred technique (Median = 1, IQR = 1 to 2), with 12 participants ranking it first and 7 ranking it second. Eye also received high preference (Median = 2, IQR = 1 to 2), with 8 participants ranking it first. Arm was typically mid-ranked (Median = 3, IQR = 3 to 4.25), with 10 participants placing it third, whereas 5 ranked it last, leading to its higher variability. In contrast to Arm, Wrist showed more stable opinions (Median = 4, IQR = 3 to 4), with most participants consistently ranking it fourth (12 in total). Finally, Finger was consistently rated as the least preferred technique (Median = 5, IQR = 4 to 5), with 13 participants placing it in the fifth position.

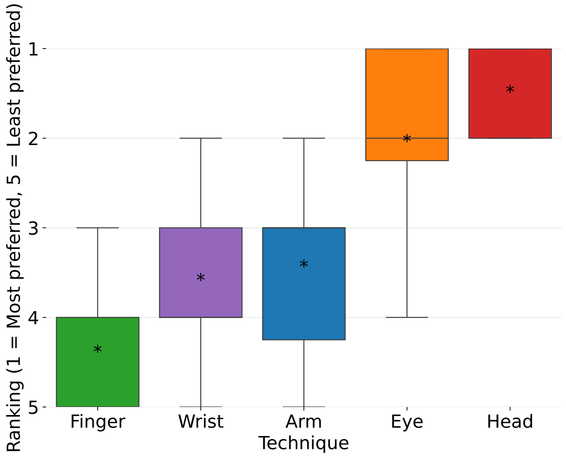

Figure 10: Subjective preference rankings per Technique.

6 DISCUSSION

We analyze our findings in terms of both performance and user experience, focusing on differences across pointing techniques and the effects of depth on selection tasks.

6.1 Pointing Techniques

Our results reveal that head- and eye-based pointing outperformed all hand-based pointing techniques across all dependent variables. Head was 40% faster than Arm, 41% faster than Wrist, and 51% faster than Finger. Eye was 35% faster than Arm, 36% faster than Wrist, and 47% faster than Finger, while Arm and Wrist were 19% and

17% faster than Finger, respectively. The comparable performance between Head and Eye techniques can be due to the coordination of eye and head movements in target acquisition [68]. Users naturally look at a target before or while turning their head towards it, making the eye- and head-based pointing faster compared to the arm, wrist, and finger. While eye-gaze offers faster initial target acquisition [38], its advantage is often offset by higher error rates, particularly at greater depths. Head had significantly lower error rates than Eye, consistent with prior work [24, 38, 49, 61]. Participants described Head as "stable" and "accurate". Three rated it as the best or most consistent technique. Among hand-based pointing techniques, Arm and Wrist performed similarly, while Finger yielded the worst results. The similar performance of Arm and Wrist can be due to them being physically linked, so arm motion also shifts the wrist pose [57]. Results based on interaction effects show more stable error rates with varying target width: Finger in small targets, Wrist in large targets, and Arm in both. This challenges previous assumptions suggesting that Finger technique could offer fast and intuitive control [14, 48]. Four participants described Finger as hard to control and difficult for selection. One stated, "Finger [was] difficult to make precise selections." NASA TLX matched performance trends: Head was rated best overall, while Finger was most demanding and frustrating.

6.2 Effect of Depth

Depth degraded performance for all techniques except Head, increasing movement time and error rates. This is consistent with perspective effects and reduced angular precision at larger distances [20, 36, 37]. However, some participants noted that Head was sometimes fatiguing, particularly for the target depth 2m, as one stated, "close targets [were] hard on [the] neck but [Head had] high accuracy." Eye was fastest at close range [59, 68], but its movement time and error rate worsened with distance [77]. One participant stated "closer targets were easier, best for big close targets." Other participants reported that they found Eye was "fast and easy" , but one noted that "it was difficult to select smaller targets". The pattern may explain why Head outperformed Eye at farther depths: because target visual angle decreased with depth, gaze-based pointing may have been more sensitive than head-based pointing to the reduced apparent target size. Arm and Wrist showed similar movement times at 2 m and 6 m, suggesting a threshold before depth costs become evident. Participants found Arm to be worse at farther distances. For example, one remarked that "Arm hurts for smaller and farther targets," while another observed that "Arm was tiring and not stable for long distance." Finger degraded between 2 m and 6 m, suggesting higher sensitivity to depth. Head remained stable across depths, with limited differences in movement time and error rate, making it a strong choice for medium to long-range selection in depth-varying AR. Movement time and error rates increased with amplitude and decreased with target width, in line with Fitts' law [21]. But some depth interactions plateaued (e.g., no significant difference of movement time between 2 m and 6 m at 2.5–3.5 m amplitudes, and no significant difference of error rate for 0.3 m vs 0.4 m widths at depths 6 m and 10 m). Throughput increased from 1.5 m to 2.5–3.5 m amplitudes, and with larger target widths. Interestingly, 0.2 m targets had significantly lower TP than wider ones, suggesting users may trade speed for precision with narrow targets. In summary, our design space analysis reveals a clear trade-off: Head is the most stable and depth-robust [40]; Eye is the fastest but depth-sensitive [38]; and Hand techniques, while familiar, suffer from fatigue and instability [53].

7 LIMITATIONS AND FUTURE WORK

We conducted the experiment under controlled stationary conditions, using a standardized 2D selection task (using ISO 9241-411) in AR. Participants reported that the solid black background improved focus by removing distractions (P18). While this controlled design enabled fair comparison, future work should examine how locomotion, lighting, and complex 3D AR environments affect performance and workload for these pointing techniques. The absence of a hand-held controller baseline is a limitation. While our aim was to compare controller-free techniques in isolation, including a standard controller-based pointing condition would have provided a useful reference for interpreting their relative performance, which future studies could explore. Another limitation is that target widths were kept constant in meters across depths, which caused visual angle to decrease as depth increased. Although this reflects many real-world AR scenarios, the reduced apparent target size at farther depths may have contributed to poorer performance for some techniques, particularly Eye, on smaller targets. Finally, our findings were derived using a Meta Quest Pro headset, which supports video see-through augmented reality. Earlier work reported differences between AR and VR [32, 54], while some more recent studies found smaller or non-significant differences between the two platforms [6]. Since our task used controller-free selection on 2D panels, we believe the relative differences between techniques found in our environment (video see-through AR) may also apply to VR and optical see-through AR. However, future work should compare video see-through AR, optical see-through AR, and VR directly to confirm the generalizability of our results.

8 CONCLUSION

We presented a systematic evaluation of five controller-free pointing techniques (Finger, Wrist, Arm, Eye, and Head) for selecting 2D targets at near (2 m), medium (6 m), and far (10 m) depths in an AR environment. By leveraging the Fitts' law, we investigated the impact of target depth along with standard Fitts' law factors (i.e., target width and amplitude) on user performance, focusing on key metrics such as movement time and error rate. Our findings indicate that head-based pointing consistently offers the optimal balance of speed and accuracy across all tested depths, outperforming arm-, wrist-, and finger-based approaches. Eye-based pointing demonstrated reaching the target fastest, but its accuracy was lower compared to head-based pointing. These results suggest that head-based pointing may serve as an effective default technique for a wide range of AR interactions, particularly when precise selection is critical across different depths. Conversely, eye-based pointing could excel in scenarios where rapid target identification is essential but tolerates slight reductions in accuracy. By identifying the strengths and weaknesses of each controller-free pointing technique, our study offers practical guidance to designers and developers of AR systems and contributes to the growing body of knowledge in human-computer interaction.

REFERENCES

[1] 2012. ISO/TS 9241-411:2012 Ergonomics of human-system interaction Part 411: Evaluation methods for the design of physical input devices. https://www.iso.org/standard/54106.html https://www.iso.org/standard/54106.html.

[2] Mohammadreza Amini, Wolfgang Stuerzlinger, Robert J Teather, and Anil Ufuk Batmaz. 2025. A Systematic Review of Fitts' Law in 3D Extended Reality. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–25. https://doi.org/10.1145/3706598.3713623

[3] Ferran Argelaguet and Carlos Andujar. 2009. Efficient 3D Pointing Selection in Cluttered Virtual Environments. *IEEE Computer Graphics and Applications* 29, 6 (Nov. 2009), 34–43. https://doi.org/10.1109/MCG.2009.117 Conference Name: IEEE Computer Graphics and Applications.

[4] Mayra Donaji Barrera Machuca and Wolfgang Stuerzlinger. 2019. The Effect of Stereo Display Deficiencies on Virtual Hand Pointing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300437

[5] Mohammad Raihanul Bashar, Mayra Donaji Barrera Machuca, Wolfgang Stuerzlinger, and Anil Ufuk Batmaz. 2025. The effect of visual depth on the vergence–accommodation conflict on 3D selection performance within virtual reality headsets: MR Bashar et al. *The Visual Computer* 41, 12 (2025), 9645–9661. https://doi.org/10.1007/s00371-025-03990-x

[6] 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 ISSN: 2642-5254.

[7] Ana M. Bernardos, Gómez , David, , and José R. Casar. 2016. A Comparison of Head Pose and Deictic Pointing Interaction Methods for Smart Environments. *International Journal of Human–Computer Interaction* 32, 4 (April 2016), 325–351. https://doi.org/10.1080/10447318.2016.1142054 Publisher: Taylor & Francis.

[8] Jonas Blattgerste, Patrick Renner, and Thies Pfeiffer. 2018. Advantages of eye-gaze over head-gaze-based selection in virtual and augmented reality under varying field of views. In *Proceedings of the Workshop on Communication by Gaze Interaction (COGAIN '18)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3206343.3206349

[9] Doug Bowman, Vinh Ly, Joshua Campbell, and Virginia Tech. 2001. Pinch keyboard: Natural text input for immersive virtual environments. (Jan. 2001). http://hdl.handle.net/10919/20010

[10] 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 (VRST '99)*. Association for Computing Machinery, New York, NY, USA, 26–33. https://doi.org/10.1145/323663.323667

[11] Gavin Buckingham. 2021. Hand Tracking for Immersive Virtual Reality: Opportunities and Challenges. *Frontiers in Virtual Reality* 2 (Oct. 2021). https://doi.org/10.3389/frvir.2021.728461 Publisher: Frontiers.

[12] Lizhou Cao, Huadong Zhang, Chao Peng, and Jeffrey T. Hansberger. 2023. Real-time multimodal interaction in virtual reality - a case study with a large virtual interface. *Multimedia Tools Appl.* 82, 16 (Feb. 2023), 25427–25448. https://doi.org/10.1007/s11042-023-14381-6

[13] Sohan Chowdhury, William Delamare, Pourang Irani, and Khalad Hasan. 2023. PAWS: Personalized Arm and Wrist Movements With Sensitivity Mappings for Controller-Free Locomotion in Virtual Reality. *Proc. ACM Hum.-Comput. Interact.* 7, MHCI (Sept. 2023), 217:1–217:21. https://doi.org/10.1145/3604264

[14] Sohan Chowdhury, A K M Amanat Ullah, Nathan Bruce Pelmore, Pourang Irani, and Khalad Hasan. 2022. WriArm: Leveraging Wrist Movement to Design Wrist+Arm Based Teleportation in VR. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 317–325. https://doi.org/10.1109/ISMAR55827.2022.00047 ISSN: 1554-7868.

[15] 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 (July 2020), 102413. https://doi.org/10.1016/j.ijhcs.2020.102413

[16] Nathan Cournia, John D. Smith, and Andrew T. Duchowski. 2003. Gaze- vs. hand-based pointing in virtual environments. In *CHI '03 extended abstracts on Human factors in computing systems - CHI '03*. ACM Press, Ft. Lauderdale, Florida, USA, 772. https://doi.org/10.1145/765891.765982

[17] E.R.F.W. Crossman. 1957. The speed and accuracy of hand movements. The Nature and Acquisition of Industrial Skill..

[18] M. Czerwinski, Greg Smith, Tim Regan, B. Meyers, G. Robertson, and G. Starkweather. 2003. Toward Characterizing the Productivity Benefits of Very Large Displays. https://www.semanticscholar.org/paper/Toward-Characterizing-the-Productivity-Benefits-of-Czerwinski-Smith/d3076f1e3c19bcf2a46922294aa67ddb1d36b34b

[19] Lisa A. Elkin, Matthew Kay, James J. Higgins, and Jacob O. Wobbrock. 2021. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. In *The 34th Annual ACM Symposium on User Interface Software and Technology (UIST '21)*. Association for Computing Machinery, New York, NY, USA, 754–768. https://doi.org/10.1145/3472749.3474784

[20] Ajoy S. Fernandes, Murdison , T. Scott, , and Michael J. Proulx. 2025. Looking in Depth: Targeting by Eye and Controller Input for Multi-Depth Target Placement. *International Journal of Human–Computer Interaction* 41, 13 (July 2025). https://doi.org/10.1080/10447318.2024.2401657 Publisher: Taylor & Francis, pages = 7952–7967,.

[21] Paul M. Fitts. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology* 47, 6 (1954), 381–391. https://doi.org/10.1037/h0055392 Place: US Publisher: American Psychological Association.

[22] Anton Franzluebbers and Kyle Johnsen. 2023. Versatile Mixed-method Locomotion under Free-hand and Controller-based Virtual Reality Interfaces. In *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology (VRST '23)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3611659.3615701

[23] Julien Gori, Olivier Rioul, Yves Guiard, and Michel Beaudouin-Lafon. 2018. The perils of confounding factors: how Fitts' law experiments can lead to false conclusions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–10. https://doi.org/10.1145/3173574.3173770

[24] John Paulin Hansen, Vijay Rajanna, I. Scott MacKenzie, and Per Bækgaard. 2018. A Fitts' law study of click and dwell interaction by gaze, head and mouse with a head-mounted display. In *Proceedings of the Workshop on Communication by Gaze Interaction (COGAIN '18)*. Association for Computing Machinery, New York, NY, USA, 1–5. https://doi.org/10.1145/3206343.3206344

[25] Faizan Haque, Mathieu Nancel, and Daniel Vogel. 2015. Myopoint: Pointing and Clicking Using Forearm Mounted Electromyography and Inertial Motion Sensors. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3653–3656. https://doi.org/10.1145/2702123.2702133

[26] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (Oct. 2006), 904–908. https://doi.org/10.1177/154193120605000909 Publisher: SAGE Publications Inc.

[27] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1063–1072. https://doi.org/10.1145/2556288.2557130

[28] David M. Hoffman, Ahna R. Girshick, Kurt Akeley, and Martin S. Banks. 2008. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision* 8, 3 (March 2008), 33. https://doi.org/10.1167/8.3.33

[29] Juan Pablo Hourcade and Natasha Bullock-Rest. 2012. How small can you go? analyzing the effect of visual angle in pointing tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. Association for Computing Machinery, New York, NY, USA, 213–216. https://doi.org/10.1145/2207676.2207706

[30] Rachel Huang, Carisa Harris-adamson, Dan Odell, and David Rempel. 2019. Design of finger gestures for locomotion in virtual reality. *Virtual Reality & Intelligent Hardware* 1, 1 (Feb. 2019), 1–9. https://doi.org/10.3724/SP.J.2096-5796.2018.0007

[31] Izabelle Janzen, Vasanth K. Rajendran, and Kellogg S. Booth. 2016. Modeling the Impact of Depth on Pointing Performance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 188–199. https://doi.org/10.1145/2858036.2858244

[32] J. Adam Jones, J. Edward Swan, Gurjot Singh, Eric Kolstad, and Stephen R. Ellis. 2008. The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization (APGV '08)*. Association for Computing Machinery, New York, NY, USA, 9–14. https://doi.org/10.1145/1394281.1394283

[33] Chaowanan Khundam. 2015. First person movement control with palm normal and hand gesture interaction in virtual reality. In *2015 12th International Joint Conference on Computer Science and Software Engineering (JCSSE)*. 325–330. https://doi.org/10.1109/JCSSE.2015.7219818

[34] Daehwa Kim, Keunwoo Park, and Geehyuk Lee. 2021. AtaTouch: Robust Finger Pinch Detection for a VR Controller Using RF Return Loss. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3411764.3445442

[35] 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 *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3706599.3719893

[36] Regis Kopper, Doug A. Bowman, Mara G. Silva, and Ryan P. McMahan. 2010. A human motor behavior model for distal pointing tasks. *International Journal of Human-Computer Studies* 68, 10 (Oct. 2010), 603–615. https://doi.org/10.1016/j.ijhcs.2010.05.001

[37] A. J. Kovacs, J. J. Buchanan, and C. H. Shea. 2008. Perceptual influences on Fitts' law. *Experimental Brain Research* 190, 1 (Sept. 2008), 99–103. https://doi.org/10.1007/s00221-008-1497-3

[38] Mikko Kytö, Barrett Ens, Thammathip Piumsomboon, Gun A. Lee, and Mark Billinghurst. 2018. Pinpointing: Precise Head- and Eye-Based Target Selection for Augmented Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173655

[39] Joseph J. LaViola, Ernst Kruijff, Ryan P. McMahan, Doug A. Bowman, and Ivan Poupyrev. 2017. *3D user interfaces: theory and practice* (second edition ed.). Addison-Wesley.

[40] Jihyeon Lee, Jinwook Kim, and Jeongmi Lee. 2023. Comparison of Virtual Reality Teleportation Targeting Method Performance depending on the Teleport Distance. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 742–745. https://doi.org/10.1109/ISMAR-Adjunct60411.2023.00160 ISSN: 2771-1110.

[41] Chiuhsiang Joe Lin, Sui-Hua Ho, and Yan-Jyun Chen. 2015. An investigation of pointing postures in a 3D stereoscopic environment. *Applied Ergonomics* 48 (May 2015), 154–163. https://doi.org/10.1016/j.apergo.2014.12.001

[42] Lars Lischke, Valentin Schwind, Kai Friedrich, Albrecht Schmidt, and Niels Henze. 2016. MAGIC-Pointing on Large High-Resolution Displays. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 1706–1712. https://doi.org/10.1145/2851581.2892479

[43] Mathias N. Lystbæk, Peter Rosenberg, Ken Pfeuffer, Jens Emil Grønbæk, and Hans Gellersen. 2022. Gaze-Hand Alignment: Combining Eye Gaze and Mid-Air Pointing for Interacting with Menus in Augmented Reality. *Proc. ACM Hum.-Comput. Interact.* 6, ETRA (May 2022), 145:1–145:18. https://doi.org/10.1145/3530886

[44] I. Scott MacKenzie. 1992. Fitts' Law as a Research and Design Tool in Human-Computer Interaction. *Human–Computer Interaction* 7, 1 (March 1992), 91–139. https://doi.org/10.1207/s15327051hci0701_3 Publisher: Taylor & Francis.

[45] Pavel Manakhov, Ludwig Sidenmark, Ken Pfeuffer, and Hans Gellersen. 2024. Filtering on the Go: Effect of Filters on Gaze Pointing Accuracy During Physical Locomotion in Extended Reality. *IEEE Transactions on Visualization and Computer Graphics* 30, 11 (Nov. 2024), 7234–7244. https://doi.org/10.1109/TVCG.2024.3456153

[46] Madhur Mangalam, Sanjay Oruganti, Gavin Buckingham, and Christoph W. Borst. 2024. Enhancing hand-object interactions in virtual reality for precision manual tasks. *Virtual Reality* 28, 4 (Nov. 2024), 166. https://doi.org/10.1007/s10055-024-01055-3

[47] Meta. 2025. Meta Quest MR, VR Headsets & Accessories. https://www.meta.com/ca/quest

[48] Domenick M. Mifsud, Adam S. Williams, Francisco Ortega, and Robert J. Teather. 2022. Augmented Reality Fitts' Law Input Comparison Between Touchpad, Pointing Gesture, and Raycast. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 590–591. https://doi.org/10.1109/VRW55335.2022.00146

[49] Katsumi Minakata, John Paulin Hansen, I. Scott MacKenzie, Per Bækgaard, and Vijay Rajanna. 2019. Pointing by gaze, head, and foot in a head-mounted display. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications (ETRA '19)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3317956.3318150

[50] Mark R. Mine. 1995. *Virtual Environment Interaction Techniques.* Technical Report. University of North Carolina at Chapel Hill, USA. https://www.cs.unc.edu/techreports/95-018.pdf

[51] Pedro Monteiro, Guilherme Gonçalves, Bruno Peixoto, Miguel Melo, and Maximino Bessa. 2023. Evaluation of Hands-Free VR Interaction Methods During a Fitts' Task: Efficiency and Effectiveness. *IEEE Access* 11 (2023), 70898–70911. https://doi.org/10.1109/ACCESS.2023.3293057 Conference Name: IEEE Access.

[52] Aunnoy K Mutasim, Anil Ufuk Batmaz, and Wolfgang Stuerzlinger. 2021. Pinch, Click, or Dwell: Comparing Different Selection Techniques for Eye-Gaze-Based Pointing in Virtual Reality. In *ACM Symposium on Eye Tracking Research and Applications (ETRA '21 Short Papers)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3448018.3457998

[53] Brad A. Myers, Rishi Bhatnagar, Jeffrey Nichols, Choon Hong Peck, Dave Kong, Robert Miller, and A. Chris Long. 2002. Interacting at a distance: measuring the performance of laser pointers and other devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '02)*. Association for Computing Machinery, New York, NY, USA, 33–40. https://doi.org/10.1145/503376.503383

[54] Djallil Naceri, Ryad Chellali, Fabien Dionnet, and Simone Toma. 2010. Depth Perception Within Virtual Environments: Comparison Between two Display Technologies. *International Journal on Advances in Intelligent Systems* 3 (Jan. 2010), 51–64. https://www.researchgate.net/publication/242330470

[55] Mathieu Nancel, Olivier Chapuis, Emmanuel Pietriga, Xing-Dong Yang, Pourang P. Irani, and Michel Beaudouin-Lafon. 2013. High-precision pointing on large wall displays using small handheld devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 831–840. https://doi.org/10.1145/2470654.2470773

[56] Bakdauren Narbayev, A K M Amanat Ullah, Jaisie Sin, Patricia Lasserre, and Khalad Hasan. 2025. Exploring Pointing and Confirmation Techniques for Teleportation Across Varying Elevations in Virtual Reality . In *2025 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE Computer Society, Los Alamitos, CA, USA, 1671–1681. https://doi.org/10.1109/ISMAR67309.2025.00170

[57] Ian Oakley, John Sunwoo, and Il-Yeon Cho. 2008. Pointing with fingers, hands and arms for wearable computing. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems (CHI EA '08)*. Association for Computing Machinery, New York, NY, USA, 3255–3260. https://doi.org/10.1145/1358628.1358840

[58] Natalia Ocampo, J Felipe Gonzalez, and Robert J Teather. 2025. Comparing Hand and Controller Avatars with Hand Tracking and Controller-Based Interaction. In *2025 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 164–174. https://doi.org/10.1109/ISMAR67309.2025.00029

[59] Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, and Hans Gellersen. 2017. Gaze + pinch interaction in virtual reality. In *Proceedings of the 5th Symposium on Spatial User Interaction (SUI '17)*. Association for Computing Machinery, New York, NY, USA, 99–108. https://doi.org/10.1145/3131277.3132180

[60] Aniruddha Prithul, Jiwan Bhandari, Walker Spurgeon, and Eelke Folmer. 2022. Evaluation of Hands-free Teleportation in VR. In *Proceedings of the 2022 ACM Symposium on Spatial User Interaction (SUI '22)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3565970.3567683

[61] Yuan Yuan Qian and Robert J. Teather. 2017. The eyes don't have it: an empirical comparison of head-based and eye-based selection in virtual reality. In *Proceedings of the 5th Symposium on Spatial User Interaction (SUI '17)*. Association for Computing Machinery, New York, NY, USA, 91–98. https://doi.org/10.1145/3131277.3132182

[62] Alexander Schäfer, Gerd Reis, and Didier Stricker. 2021. Controlling Teleportation-Based Locomotion in Virtual Reality with Hand Gestures: A Comparative Evaluation of Two-Handed and One-Handed Techniques. *Electronics* 10, 6 (Jan. 2021), 715. https://doi.org/10.3390/electronics10060715 Number: 6 Publisher: Multidisciplinary Digital Publishing Institute.

[63] Alexander Schäfer, Gerd Reis, and Didier Stricker. 2022. Controlling Continuous Locomotion in Virtual Reality with Bare Hands Using Hand Gestures. In *Virtual Reality and Mixed Reality*, Gabriel Zachmann, Mariano Alcañiz Raya, Patrick Bourdot, Maud Marchal, Jeanine Stefanucci, and Xubo Yang (Eds.). Springer International Publishing, Cham, 191–205. https://doi.org/10.1007/978-3-031-16234-3_11

[64] Mengdi Shi, Tao Hu, and Jiawen Yu. 2022. Pointing Cursor Interaction in Virtual Reality from the Perspective of Distance Perception | IIETA. https://doi.org/10.18280/ts.390209

[65] Garth Shoemaker, Takayuki Tsukitani, Yoshifumi Kitamura, and Kellogg S. Booth. 2012. Two-Part Models Capture the Impact of Gain on Pointing Performance. *ACM Trans. Comput.-Hum. Interact.* 19, 4 (2012), 28:1–28:34. https://doi.org/10.1145/2395131.2395135

[66] Shaishav Siddhpuria, Sylvain Malacria, Mathieu Nancel, and Edward Lank. 2018. Pointing at a Distance with Everyday Smart Devices. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3173574.3173747

[67] Ludwig Sidenmark and Hans Gellersen. 2019. Eye, Head and Torso Coordination During Gaze Shifts in Virtual Reality. *ACM Trans. Comput.-Hum. Interact.* 27, 1 (2019), 4:1–4:40. https://doi.org/10.1145/3361218

[68] Ludwig Sidenmark and Hans Gellersen. 2019. Eye&Head: Synergetic Eye and Head Movement for Gaze Pointing and Selection. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. Association for Computing Machinery, New York, NY, USA, 1161–1174. https://doi.org/10.1145/3332165.3347921

[69] Ludwig Sidenmark, Franziska Prummer, Joshua Newn, and Hans Gellersen. 2023. Comparing Gaze, Head and Controller Selection of Dynamically Revealed Targets in Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics* 29, 11 (Nov. 2023), 4740–4750. https://doi.org/10.1109/TVCG.2023.3320235

[70] Siddhanth Raja Sindhupathiraja, A K M Amanat Ullah, William Delamare, and Khalad Hasan. 2024. Exploring Bi-Manual Teleportation in Virtual Reality. In *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. 754–764. https://doi.org/10.1109/VR58804.2024.00095 ISSN: 2642-5254.

[71] 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, 6 (Dec. 2004), 751–789. https://doi.org/10.1016/j.ijhcs.2004.09.001

[72] Da Tao, Xiaofeng Diao, Tieyan Wang, Jingya Guo, and Xingda Qu. 2021. Freehand interaction with large displays: Effects of body posture, interaction distance and target size on task performance, perceived usability and workload. *Applied Ergonomics* 93 (May 2021), 103370. https://doi.org/10.1016/j.apergo.2021.103370

[73] Robert J. Teather, Andriy Pavlovych, Wolfgang Stuerzlinger, and I. Scott MacKenzie. 2009. Effects of tracking technology, latency, and spatial jitter on object movement. In *2009 IEEE Symposium on 3D User Interfaces*. 43–50. https://doi.org/10.1109/3DUI.2009.4811204

[74] Robert J. Teather and Wolfgang Stuerzlinger. 2013. Pointing at 3d target projections with one-eyed and stereo cursors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 159–168. https://doi.org/10.1145/2470654.2470677

[75] Eleftherios Triantafyllidis and Zhibin Li. 2021. The Challenges in Modeling Human Performance in 3D Space with Fitts' Law. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3411763.3443442

[76] Uta Wagner, Andreas Asferg Jacobsen, Tiare Feuchtner, Hans Gellersen, and Ken Pfeuffer. 2024. Eye-Hand Movement of Objects in Near Space Extended Reality. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology (UIST '24)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3654777.3676446

[77] Uta Wagner, Mathias N. Lystbæk, Pavel Manakhov, Jens Emil Sloth Grønbæk, Ken Pfeuffer, and Hans Gellersen. 2023. A Fitts' Law Study of Gaze-Hand Alignment for Selection in 3D User Interfaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3544548.3581423

[78] Franziska Westermeier, Larissa Brübach, Carolin Wienrich, and Marc Erich Latoschik. 2024. Assessing Depth Perception in VR and Video See-Through AR: A Comparison on Distance Judgment, Performance, and Preference. *IEEE Transactions on Visualization and Computer Graphics* 30, 5 (May 2024), 2140–2150. https://doi.org/10.1109/TVCG.2024.3372061 Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[79] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[80] Dennis Wolf, Jan Gugenheimer, Marco Combosch, and Enrico Rukzio. 2020. Understanding the Heisenberg Effect of Spatial Interaction: A Selection Induced Error for Spatially Tracked Input Devices. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3313831.3376876

[81] Wenge Xu, Hai-Ning Liang, Anqi He, and Zifan Wang. 2019. Pointing and Selection Methods for Text Entry in Augmented Reality Head Mounted Displays. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 279–288. https://doi.org/10.1109/ISMAR.2019.00026 ISSN: 1554-7868.

[82] Difeng Yu, Tilman Dingler, Eduardo Velloso, and Jorge Goncalves. 2024. Object Selection and Manipulation in VR Headsets: Research Challenges, Solutions, and Success Measurements. *ACM Comput. Surv.* 57, 4 (2024), 98:1–98:34. https://doi.org/10.1145/3706417

[83] Shumin Zhai, Jing Kong, and Xiangshi Ren. 2004. Speed–accuracy tradeoff in Fitts' law tasks—on the equivalency of actual and nominal pointing precision. *International Journal of Human-Computer Studies* 61, 6 (Dec. 2004), 823–856. https://doi.org/10.1016/j.ijhcs.2004.09.007

[84] Fan Zhang, Shaowei Chu, Ruifang Pan, Naye Ji, and Lian Xi. 2017. Double hand-gesture interaction for walk-through in VR environment. In *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*. 539–544. https://doi.org/10.1109/ICIS.2017.7960051

[85] Caijun Zhao, Kai Way Li, and Lu Peng. 2023. Movement Time for Pointing Tasks in Real and Augmented Reality Environments. *Applied Sciences* 13, 2 (Jan. 2023), 788. https://doi.org/10.3390/app13020788 Number: 2 Publisher: Multidisciplinary Digital Publishing Institute.