Quantifying 3D Pointing: Characterizing What Happens During Pointing Selection in Virtual Reality

Junhan Kong
University of Washington
Seattle, Washington, USA
junhank@uw.edu

Jacqui Fashimpaur
Reality Labs Research
Meta Inc.
Redmond, Washington, USA
jfashimpaur@gmail.com

Michael J Proulx
Reality Labs Research
Meta Inc.
Redmond, Washington, USA
michaelproulx@meta.com

Hemant Bhaskar Surale
Reality Labs Research
Meta Inc.
Toronto, Ontario, Canada
hemant.surale@gmail.com

Jacob O. Wobbrock
University of Washington
Seattle, Washington, USA
wobbrock@uw.edu

Amy Karlson
Reality Labs Research
Meta Inc.
Redmond, Washington, USA
akkarlson@gmail.com

Abstract

Measures of target selection in virtual reality (VR) have focused mainly on overall performance, such as speed and accuracy. However, a selection is not atomic—numerous things happen *during* the short "click" when a target is selected. In this work, we formalize a set of measures that provide insight into what happens *during* this process. The measures characterize the pointer movement, user behaviors (gaze, head, hand and finger), and coordination between modalities. To validate the measures, we collected pointing data from 36 participants using five different 3D pointing techniques. Our results show that these measures effectively characterize pointing performance, task load, perceived exertion, ease of use/learning, comfort, and sense of agency. We found that tolerating less "stable" pointing and considering alignment between modalities both improve pointing performance. Our findings suggest that 3D pointing should prioritize speed, allow for more movements during selection, and consider coordination of different modalities.

CCS Concepts

• **Human-centered computing** → **Virtual reality**; **Pointing**.

Keywords

Pointing metrics, human performance, gaze input, hand gestures, eye tracking, hand tracking, virtual reality, head-mounted devices

ACM Reference Format:

Junhan Kong, Jacqui Fashimpaur, Michael J Proulx, Hemant Bhaskar Surale, Jacob O. Wobbrock, and Amy Karlson. 2026. Quantifying 3D Pointing: Characterizing What Happens During Pointing Selection in Virtual Reality. In *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION (ICMI '26), October 05–09, 2026, Napoli, Italy*. ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3776574.3831138

ACM ISBN 979-8-4007-2318-6/26/10
https://doi.org/10.1145/3776574.3831138

1 Introduction

With the rapid development of virtual reality (VR) devices, pointing techniques utilizing various input modalities have been invented, for example, using controllers [52, 60], hand gestures [12, 48], gaze [44], or on-body devices. Prior work compared pointing techniques on their performance and user preference, but often without a consistent measure for comparing these techniques [45]. Meanwhile, most existing performance measures focus on high-level and overall measures, such as speed and accuracy.

However, a 3D target selection process is not atomic [42]. Even during a short, quick "click", numerous things happen between the start and end of selection trigger, for example, changes in user's gaze point, hand posture or controller position, etc. Such characteristics are not unique to 3D pointing—prior work investigated what happens *during* selection with mice [26, 28, 39] and touch-based systems [30, 31], devising measures characterizing device and user performance. Motivated by similar ideas, we investigate what happens *during* a 3D target selection in VR. The aim is to understand how the pointer traces and user behavior during this process are related to the pointing experience. This approach has the potential to reveal causes of inaccuracies, discomfort, or fatigue, and inform improvements to these techniques.

In this work, we offer eight measures that can effectively characterize pointing performance and user experience—by looking into what happens during the short "click" of selection. The measures consist of four that characterize user behavior within a single modality: gaze direction range, head direction range, palm position drift, and index fingertip variability. Additionally, four measures assess pointer movement and the alignment between modalities: pointer direction variability, hit point variability, mean alignment distance, and alignment-adjusted endpoint. These measures were selected from a total of 31 measures investigated, based on their significance of correlation with performance and user experience ratings, conceptual clarity, and ease of interpretation.

To validate these measures, we ran a lab study to collect pointing data with 36 participants and five pointing techniques: using the controller, gaze and pinch, gaze and the controller button, handray, and a ring prototype [29]. We collected 13,885 valid pointing trials, computed measure values for each trial, and ran descriptive and inferential statistics on them. Our results show that these measures can effectively characterize differences among techniques and align with pointing throughput. We also found significant correlations between these measures and accuracy, task load, perceived exertion, ease of use/learning, comfort, and sense of agency.

The empirical findings revealed distinct patterns of user behavior during 3D pointing. First, speed had larger effects on overall performance than accuracy across all the techniques, and higher tolerance for less "stable" pointing reduces task load and exertion. Second, we found that different modalities work in coordination, and using our alignment-adjusted endpoint measure improves pointing accuracy. Our findings suggest that 3D pointing should improve tolerance for less "stable" pointing, prioritize movement speed and lower latency over pointing and sensing accuracy, and consider coordination of different input modalities. The contributions of this work are:

- Formalizing eight 3D measures that characterize pointer movement, gaze, head, and hand movements, pointer interaction with task planes, and alignment of a user's gaze and their pointer;
- Validation of these measures with five different 3D pointing techniques in a formal study of 36 participants;
- Empirical results of 3D pointing behavior and their design implications through investigating the measures with different pointing techniques.

2 Related Work

Understanding Human Pointing Behavior. Numerous works have examined human pointing behavior, particularly with traditional graphical user interfaces (GUIs). MacKenzie et al. [39] devised accuracy measures that include movement variability, offset, error and more for evaluating mouse pointing, revealing differences that enable detailed comparison of input devices. Keates et al. [28] and Hwang et al. [26] extended MacKenzie et al.'s work, characterizing pointing behavior by people with motor impairments, and Kong et al. [30, 31] devised touchscreen measures that characterize fine motor abilities. Prior work also modeled how users perceive their own touch point [24, 25], conceptualized touch selection as a probabilistic distribution around the target [6], and quantified the relative precision of touches due to the speed-accuracy tradeoff [5]. Our work takes inspiration from MacKenzie et al. [39] and Kong et al. [31] for traditional 2D mouse and touchscreens, and aims to investigate similar measures for pointing in 3D environments.

Pointing Techniques for 3D. Controller, hand, and gaze pointing are being widely used in head-mounted devices (HMD). One common technique is the *Handray*, leveraging the "virtual pointer" [46] of a ray-cast from a user's palm enabling distant object manipulation, in contrast to the "virtual hand" [46], which better supports nearby interactions. Gaze-based techniques have also been shown to be effective, both with 2D graphical UIs [49, 59] and virtual or augmented environments [32, 45, 53], for example, the "gaze and pinch" [32, 34, 44]. Novel 3D pointing techniques have been invented leveraging a combination of input modalities [1], such as integrating depth into mouse control [62], combining gaze and hand gestures [11, 37, 44, 47, 59], integrating eye and head motion [50].

Prior work also compared different 3D pointing techniques, highlighting their strengths and weaknesses. For example, prior work found gaze improving speed but weakening spatial information recall compared to finger pointing [53], faster and more preferred than mid-air cursors [54], and advantageous for interacting with distant objects [58]; gaze was also found to be slower than head- and foot-based pointing [41], had lower cognitive load and lower accuracy compared to controllers [36] and degraded performance on smaller targets [15]. Prior work also found dwelling more accurate yet significantly slower than pinch and button control [43], and faster selection and reduced hand fatigue through combining gaze with hand input [61].

Characterizing 3D Pointing Performance. To characterize 3D pointing performance, a large body of prior work applied Fitts' Law [16] with adjustments enabling transformation from 2D to 3D, for example, in investigating gaze and hand alignment during 3D menu selection [55], click-and-dwell interactions using gaze, head, and mouse [18], the effects of rotational jitter on mid-air interactions [2–4]. Different from these experiments, this work primarily focuses on the short duration of selection attempts.

Metrics have been devised to characterize physical exertion caused by 3D pointing, the most popular being consumed endurance [22]. Other arm exertion measures include Jang et al.'s [27] model for estimating active, rest, and fatigue muscle states. Prior work also observed pointing errors caused by movement during selection, namely the "Heisenberg effect" [9], which was found to be responsible for a significant proportion of pointing errors [57]. Our work expands upon this prior work to more modalities and beyond overall performance measures like accuracy.

Indeed, some similar measures have been previously used in individual studies on an *ad hoc* basis (Table 1). Yet, this work statistically validates correlations of each measure with existing high-level measures like pointing performance and task load. Furthermore, research has found inconsistencies of results in prior evaluations due to the varying tasks and ratings [45, 55], and called for the need of a "set of common tasks and evaluation metrics" [45]—which is exactly what this work is offering.

	Prior Work Leveraging Similar Measures	**Measures as Described in Prior Work**
Our Measure (1/8)		
Palm Position Drift	Yu et al. [58], Wagner et al. [55]	Hand Movement Distance, Hand Rotation Angles
Additional Measures We Investigated (4/31)		
Target Hit Ratio, Pointer Direction Drift	Wolf et al. [57]	HeisenbergError, HeisenbergMagnitude
Target Re-Entry	Luro and Sundstedt [36]	Number of times the participant lost the target during aiming
Palm Position Range	Luong et al. [35]	Hand Movement Space

Table 1: Prior work using similar measures to evaluate specific techniques, and measures as described in prior work.

3 Pointing Measures

In this section, we formalize eight measures that we use to characterize what happens between "pointer start" and "pointer release" during 3D target selection in virtual environments.

3.1 Anatomy of 3D Pointing

Our measures intend to capture movements during the "click" of pointing selection that consists of at least an active pointer and a

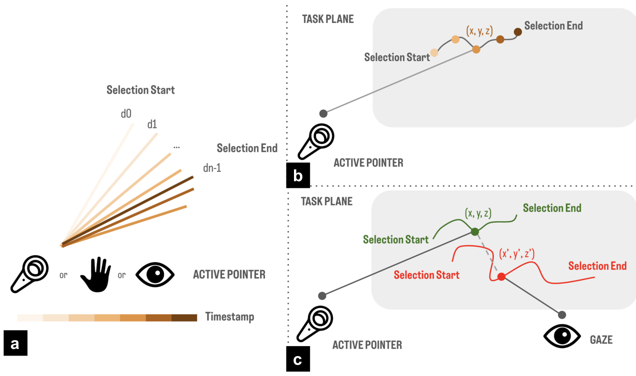

Figure 1: (a) Formalization of a pointing trace as a time series of 6 DoF ray-casts from selection-start to selection-end, using a controller, hand, or eye gaze; (b) Illustration of hit point variability; (c) Illustration of mean alignment distance.

selection trigger, similar to Wagner et al. [55]'s conceptualization. A selection trigger should consist of at least a start and end event, for example, controller button-down and button-up. A pointing process can be defined as a sequence of pointer events from selection-start at timestamp t_0 to selection-end at timestamp t_{n-1} (Figure 1). For the i^{th} pointer event: $p_i = (p_{x,i}, p_{y,i}, p_{z,i})$ is the origin position of the ray-cast; $d_i = (d_{x,i}, d_{y,i}, d_{z,i})$ is the direction of the ray-cast. To simplify presentation, we also define the angle change from d_i to d_j as $\Delta(d_i, d_j)$ and the Euclidean distance between 3D points p_i and p_j as $D(p_i, p_j)$. We define gaze, head, hand and finger movement events captured from selection-start to selection-end as follows:

- Gaze Movement: $G_0, ..., G_{n_g-1}$, where $G_i = (p_{x_g,i}, p_{y_g,i}, p_{z_g,i}, d_{x_g,i}, d_{y_g,i}, d_{z_g,i})$ is the 6 DoF gaze ray direction;
- Head Movement: $H_0, ..., H_{n_h-1}$, where $H_i = (p_{x_h,i}, p_{y_h,i}, p_{z_h,i}, d_{x_h,i}, d_{y_h,i}, d_{z_h,i})$ is the 6 DoF head ray direction;
- Palm Movement: $P_0, ..., P_{n_p-1}$, where $P_i = (p_{x_p,i}, p_{y_p,i}, p_{z_p,i})$ is the palm position;
- Index Fingertip Movement: $f_{0,\text{idx}}, ..., f_{n_f-1,\text{idx}}$, where $f_{i,\text{idx}} = (p_{x_f,i,\text{idx}}, p_{y_f,i,\text{idx}}, p_{z_f,i,\text{idx}})$ is the index fingertip position.

3.2 Pointer Measures

We formalize two measures for the active pointer movement in 3D and on the 2D task plane.

Pointer Direction Variability. Pointer direction variability is the cumulative change of pointing direction over the duration of selection. Intuitively, this measure indicates how "jittery" the pointing direction was during a selection.

$$\text{Pointer Direction Variability} = \sum_{i=1}^{n-1} \Delta(d_{i-1}, d_i) \in [0, \infty)$$

Hit Point Variability. For the i^{th} pointer event, we define $\text{Hit}_i = (x_{\text{hit},i}, y_{\text{hit},i})$ as the position where the pointer ray-cast hits the task plane, in the plane's 2D coordinates (Figure 1b). Hit point variability is the total distance of successive hit point movements from selection-start to selection-end. Intuitively, this measure indicates how "jittery" the pointer movement is on the task plane.

$$\text{Hit Point Variability} = \sum_{i=1}^{n-1} D(\text{Hit}_{i-1}, \text{Hit}_i) \in [0, \infty)$$

3.3 Single Modality Measures

We formalize four measures for movement in each modality.

Gaze Direction Range. Gaze direction range is the angle between the two gaze directions that are farthest apart.

$$\text{Gaze Direction Range} = \left(\max_{i,j \in 0,\ldots,n_g-1} \Delta(d_{g,i}, d_{g,j}) \right) \in [0, \infty)$$

Head Direction Range. Head direction range is the angle between the two head directions that are farthest apart.

$$\text{Head Direction Range} = \left(\max_{i,j \in 0,\ldots,n_h-1} \Delta(d_{h,i}, d_{h,j}) \right) \in [0, \infty)$$

Palm Position Drift. Palm position drift is the Euclidean distance between the palm positions at selection-start and selection-end, indicating how far the user's hand moved from where it started.

$$\text{Palm Position Drift} = D(P_0, P_{n_p-1}) \in [0, \infty)$$

Index Fingertip Variability. Index fingertip variability is the total distance covered by successive index fingertip movements from selection-start to selection-end.

$$\text{Index Fingertip Variability} = \sum_{i=1}^{n_f-1} D(f_{i-1,\text{idx}}, f_{i,\text{idx}}) \in [0, \infty)$$

3.4 Alignment Measures

We formalize two measures for the alignment between modalities.

Mean Alignment Distance. Alignment distance is the average distance between the active pointer hit point $\text{Hit}_i = (x_{\text{hit},i}, y_{\text{hit},i})$ and the gaze hit point $\text{Hit}'_i = (x'_{\text{hit},i}, y'_{\text{hit},i})$ on the task plane. Figure 1c illustrates an example of alignment distance during pointing.

$$\text{Mean Alignment Distance} = \frac{1}{n} \sum_{i=0}^{n-1} D(\text{Hit}_i, \text{Hit}'_i) \in [0, \infty)$$

Alignment-Adjusted Endpoint. Alignment-adjusted endpoint is the 3D pointer position at the time when the active pointer and gaze hit point locations are closest on the task plane. Intuitively, this measure offers the pointer position when the active pointer and the user eye gaze are the most "aligned".

$$\text{Alignment-Adjusted Endpoint} =$$
$$(x_{i^*}, y_{i^*}, z_{i^*}) \text{ where } i^* = \operatorname*{argmin}_{i=0,\ldots,n-1} D(\text{Hit}_i, \text{Hit}'_i)$$

In this work, we investigated a total of 31 measures, and selected eight measures based on two criteria: (1) correlation with performance and user experience ratings, and (2) conceptual clarity and ease of interpretation. We first selected measures based on their significance of correlations in our analyses with the full set of 31 measures (Appendix E); among measures exhibiting similar statistical significance, we then selected one measure for each modality to avoid repetition yet retain high interpretability. For example, we chose range of motion for gaze and head movements while using finer-grained measures for pointer movements, and we chose palm movements over average joint movements. We include formalizations of all 31 measures and their distributions and analyses in Appendices A-E.

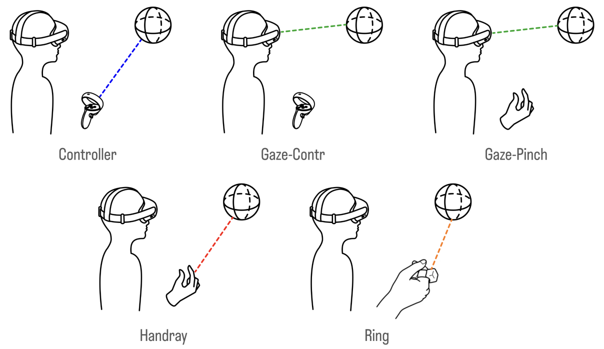

Figure 2: The five techniques evaluated: *Controller*, *Gaze-Contr*, *Gaze-Pinch*, *Handray*, and *Ring*. The dashed lines illustrate the ray-cast used for pointing in each technique.

4 Study Method

We conducted a controlled lab study with 36 participants to assess how well our measures characterize 3D pointing behaviors. Unlike prior work that primarily compares techniques, we also investigated the correlations between our measures and existing pointing **performance measures** (speed, accuracy, throughput) and **user experience ratings** (task load, perceived exertion, ease of use/learning, comfort, sense of agency).

4.1 Experiment Conditions

4.1.1 Pointing Techniques. We chose five pointing techniques:

- ***Controller***. The user points with the default controller ray and selects targets using the controller button;
- ***Gaze-Contr***. The user points with their gaze and selects targets using the controller button;
- ***Gaze-Pinch***. The user points with their gaze and selects using a pinch gesture with the thumb and the index finger;
- ***Handray***. The user points with a ray-cast emanating from their palm and selects using a pinch gesture with the thumb and the index finger;
- ***Ring***. The user points with a ray-cast emanating from the ring position (Figure 2) and selects by tapping their thumb on the middle phalanx of the index finger (Figure 4b).

These techniques were selected to cover a range of different modalities (gaze, hand, finger, controller) and selection triggers (pinch, thumb-tap, button). They also enable pairwise comparisons that isolate either the pointing modality or the selection trigger. Furthermore, the techniques range from widely adopted controller-based interaction to novel wearables, supporting assessment of the measures' applicability across various devices and techniques.

4.1.2 Target Configurations. We used spherical targets with two **target size** conditions: small and large (5° and 10° diameters), following prior work [14, 55]. Other factors (e.g., depth, shape) were omitted due to fatigue observed in a pilot study including both size and depth. As full counterbalancing was infeasible for a within-subjects study with 36 participants and 10 conditions (5 *Technique* × 2 *Target Size*), we used constrained randomization to generate 36 unique presentation orders, ensuring balanced distribution across ordinal positions. Each technique appeared 5–9 times per position, within ±2 of the theoretical mean.

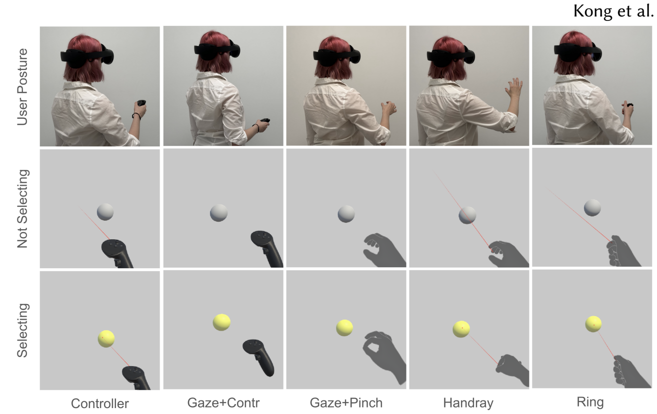

Figure 3: The five pointing techniques: user posture during pointing (top), default state without target hit or selection (middle), and state during target selection (bottom).

4.2 Existing Pointing Task Measures

We validate our measures by comparing them to existing pointing task measures as below:

- ***Hit or Miss***. We record whether each trial was a "hit," meaning that the pointing ray-cast collides with the target sphere at both selection-start and selection-end;
- ***Task Load***. We measure task load using the NASA-TLX questionnaire [19, 20]. We record user self-reported physical demand, mental demand, temporal demand, performance, effort, and frustration, each from 0 to 10;
- ***Perceived Exertion***. We measure perceived exertion using the Borg CR-10 scale [8]. We record the user's overall perceived exertion as well as exertion in the eye, head and neck, hand and finger, and arm, each from 0 to 10;
- ***Additional User Experience Ratings***. We include additional measures taking inspiration from prior work [35], including ease of use, ease of learning, overall comfort, and sense of agency, each from 0 to 10.

4.3 Pointing Testbed

We developed a Unity testbed app that generates pointing targets and records user behavior data. The application ran in Unity version 2022.3.38f1, on a high-end Windows 10 machine (2.7GHz Intel i7 CPU, 1.73GHz GeForce GTX 3070 GPU).

4.3.1 Eye Calibration. We used the Meta Quest Pro default eye calibration, and validated calibration results through accuracy and precision of gaze on 10 blinking cross-hair targets on a circle of 10 degrees radius around the center of the view.

4.3.2 Pointing Task. For each trial, the testbed generates a gray target sphere (Figure 3) at a random position within ±35° horizontally and vertically from the center of view, at a fixed distance of 2 meters, a comfortable range within users' view and reach [10]. Upon selection, a sound is played and the next target is generated until the block ends. The testbed provides a yellow hover effect and visualizes pointer ray-casts as red beams with a red cursor at the hit point. Ray visualization is enabled for controller, hand, and ring, but not gaze, following prior work [14], as it improves pointing accuracy for hand/controller but can degrade gaze performance due to cursor chasing [40].

4.3.3 Pointing Techniques and User Behavior Tracking. We used a Meta Quest Pro headset (1800×1920 px per eye, 90 Hz refresh rate, $106° \times 96°$ FoV) and a ring prototype (the ElectroRing [29]). *Gaze-Contr* and *Gaze-Pinch* use the Meta OVRPlugin library for gaze ray-casting, while *Handray* and *Gaze-Pinch* use the OVRHand library for hand ray-casting and pinch detection. For the *Ring* technique, the ring (Figure 4a–b) is worn on the dominant index finger. The ray-cast origin is computed from hand joints (OVRSkeleton) as the midpoint of the index proximal phalanx, and its direction from accumulated gyroscope motion, calibrated to a forward-facing baseline (Figure 4c). Selection is triggered via thumb tap, detected through electrode voltage changes [29]. The controller ray-cast information was recorded at a 72 Hz sampling rate. Selection-start and selection-end events with timestamps were also recorded. The hand pointer 6 DoF ray-cast information was obtained through the Meta OVRHand library and recorded at a 60 Hz sampling rate. Gaze and head position and direction were obtained through the Meta OVRPlugin library and recorded at a 72 Hz sampling rate.

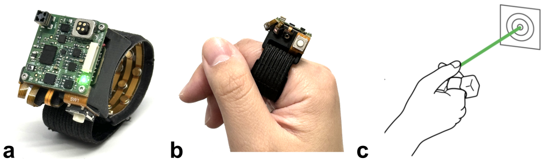

Figure 4: (a) *Ring* hardware (ElectroRing [29]) with a 9-axis IMU and embedded electrodes; (b) "thumb tap" gesture for selection; (c) the ring pointing posture used during interaction and calibration.

4.4 Participants

We recruited 36 participants using email solicitations. Participants were 36.1 years old on average ($SD = 12.6$), including 14 women and 22 men. Thirty-five participants were right-handed and one participant was left-handed. On a scale of 0 to 10, participants self-reported having moderate ($M = 4.8, SD = 2.6$) experience with XR overall; for pointing techniques, participants self-reported having moderate ($M = 5.7, SD = 2.9$) experience using controllers, moderate ($M = 4.3, SD = 3.3$) experience using freehand gestures, little to moderate experience ($M = 3.2, SD = 3.6$) using gaze, and no or little experience ($M = 0.5, SD = 1.7$) using ring-based devices.

4.5 Procedure

Study sessions were conducted in person with IRB approval. Each session consisted of five parts, one per technique, following the same procedure. At the start of each part, participants completed eye calibration on the headset by following blinking targets, aiming to keep tracking error within 1.5°. For the *Ring* condition, participants completed an additional calibration of pointing forward, holding posture, performing a thumb tap, and aligning to the center of the view (Figure 4c). Each *Technique* × *Target Size* condition included 5 practice and 40 test trials. After each block, participants verbally reported responses to rating questions (Section 4.2).

4.6 Data Analysis

4.6.1 Preprocessing. We removed outlier trials with (1) no more than two sensor data points or (2) excessively long duration (above $M + 2SD$, 0.93 seconds), likely due to sensor errors. We then computed the eight measures for each valid trial. No additional smoothing was applied beyond the library default to avoid eliminating useful information from short sensor streams (96.1% trials had no more than 20 sensor data points). We removed P27 from the additional ratings analysis as there was no variation in their ratings.

4.6.2 Pointing Throughput Calculations. To assess pointing *performance*, we computed throughput (bits/s) [38]. We excluded trials with movement amplitudes below $A/2$ [38], and regressed movement time (MT) on effective index of difficulty (ID_e), calculated from effective amplitude (A_e) and width (W_e) [51]. As ray-cast interactions lack 3D endpoints, we estimated W_e using the observed error rate [51]. Our study did not include predefined amplitude (A) by target width (W) conditions, so we grouped trials via *post hoc* clustering of amplitudes by median A_e to form pseudo-conditions [13]. Together with target size, this produced four conditions, and thus four (ID_e, MT) pairs, per participant–technique pair.

4.6.3 Counterbalancing Verification. The active pointer measure showed significant effects of *Order* ($\chi^2_{9,N=13885} = 132.28, p < .01$) and *Order* × *Technique* ($\chi^2_{36,N=13885} = 281.73, p < .01$). We therefore included *Order* as a covariate in comparisons across techniques. In the adjusted model, the order effect remained statistically significant ($\chi^2_{9,N=13885} = 109.84, p < .01$) but had a minimal effect size ($\eta^2_p = 0.008$), whereas the *Technique* effect was substantially larger ($\chi^2_{4,N=13885} = 20674.40, p < .01, \eta^2_p = 0.600$). This result suggests that controlling for order *revealed* rather than *masked* true technique differences. No significant order effects were found for the user experience ratings (task load, perceived exertion, ease of use/learning, comfort, sense of agency).

4.6.4 Statistical Analyses. To understand how the measure values differ across *Technique* and *Target Size*, we ran linear mixed models (LMMs) [17, 33] on each measure[1], with *Technique*, *Target Size*, and their interaction as fixed effects, *Order* as a covariate, and *Participant* as a random effect. *Trial* was excluded from the random effects as it contributed zero variance. We applied log transformations to all dependent variables except the alignment measures, and verified normality through examination of residuals. Significant omnibus tests were followed by *post hoc* pairwise comparisons with Holm's correction [23]. To investigate the correlations of our measures with pointing performance, we ran generalized linear mixed models (GLMMs) [7] on the binomial trial *Hit*, using pointer direction variability or endpoint type (with *vs.* without alignment adjustment) as fixed effects, and *Participant* and *Trial* as random effects. To investigate the correlations of our measures with user experience, we used mixed ordinal logistic regression [21] on NASA-TLX and additional user ratings, with each pointer measure[2] as a fixed effect and *Participant* as a random effect. Borg CR-10 scores were analyzed using LMMs with each pointer measure as a fixed effect and *Participant* as a random effect[3].

[1]As the four target hit measures among the full set of 31 measures capture higher-level outcomes rather than movement traces, we report descriptive statistics but exclude them from inferential analyses.
[2]Measures were averaged per *Technique* × *Target Size* block, as ratings were collected once per block.
[3]Borg CR-10 ratings were treated as continuous as they contained 0.5 values.

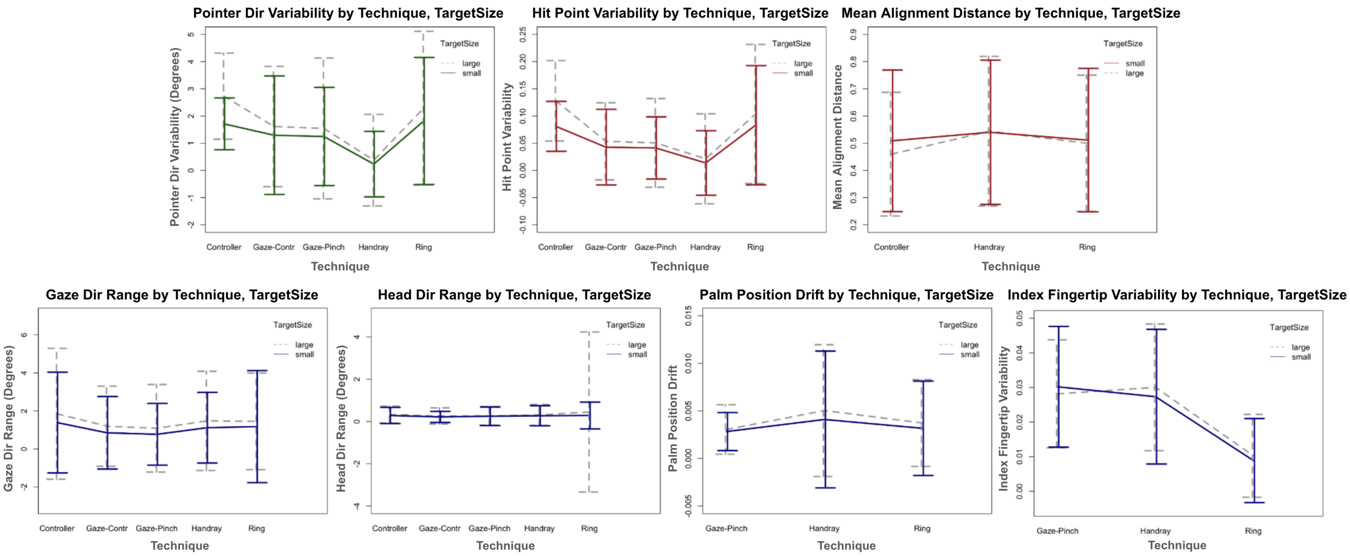

Figure 5: Interaction plots ($M \pm SD$) showing gaze direction range, head direction range, palm position drift, index fingertip variability, pointer direction variability, hit point variability, and mean alignment distance by *Technique* $\times$ *Target Size*.

5 Results

In this section, we present analysis results showing how the measures characterize pointing behavior. We compare measures across techniques and target sizes, then present correlations between our measures and existing performance and user experience measures.

5.1 Summary of Data

We collected 14112 trials of pointing data from the 36 participants, and there were 13885 (98.4%) valid trials remaining after outlier removal. The valid pointing trials had an average duration of 0.17 seconds ($SD = 0.08$), and contained an average of 12 data points in the pointer traces ($SD = 5$). Overall, 12427 (89.5%) trials successfully hit the target, while 1458 trials (10.5%) missed the target.

5.2 Comparing Different Pointing Techniques

Overall, *Handray* had the lowest pointer measure values across all five techniques, followed by the two gaze-based techniques, *Gaze-Contr* and *Gaze-Pinch*; *Controller* and *Ring* had slightly higher pointer measure values, indicating overall more pointer movement during a selection. *Controller* had the smallest standard deviation of pointer measures, while *Ring* had the largest standard deviation, indicating the amount of variation in user behavior. We found significant effects of *Technique*, *Target Size*, and their interaction on the measure values for most of our measures. We present the full omnibus test results and distributions of all user ratings and all measure values in Appendices D and E.

5.3 How Our Measures Characterize Pointing Performance

In this section, we explore how our measures relate to pointing accuracy and throughput.

Lower measure values (i.e., more "stable" pointing) are associated with higher pointing accuracy. We show the pointing accuracy of each *Technique* and *Target Size* in Figure 6. Among the five techniques, *Controller* had the highest pointing accuracy overall, followed by *Handray*. The distribution of pointer direction variability by accuracy (Figure 6b) also shows that pointer direction variability was negatively associated with accuracy, i.e., the more "stable" the pointing was, the higher pointing accuracy was. We found significant correlations between pointing accuracy and pointer direction variability ($\chi^2_{1,N=13885} = 538.1, p < .01$).

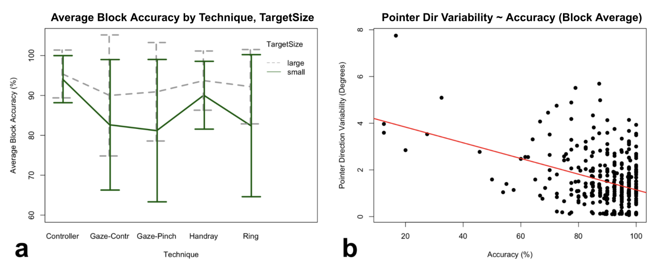

Figure 6: (a) Interaction plot showing pointing accuracy (error bar representing $M \pm SD$) of each *Technique* $\times$ *Target Size*; (b) Distribution of pointer direction variability by accuracy. The smaller the pointer measure was (i.e., the more "stable" the pointing was), the higher pointing accuracy was.

However, techniques with higher measure values (i.e. less "stable" pointing) exhibit better overall pointing performance. In our throughput analysis, we obtained a total of 13018 movement traces from the data, and there were 12988 (99.8%) valid trials remaining after outlier removal. The *Controller* exhibited the highest overall throughput ($M = 2.49, SD = 0.51$) across participants; *Gaze-Contr* ($M = 2.23, SD = 0.78$) and *Gaze-Pinch* ($M = 1.94, SD = 0.60$) exhibited slightly lower throughput, followed by *Ring* ($M = 1.75, SD = 0.50$); *Handray* exhibited the lowest throughput among all five techniques ($M = 1.34, SD = 0.27$). [4]

[4]While we did not perform a statistical analysis on the correlation between throughput and our measures due to the limited amount of throughput data available (one throughput per participant–technique pair), the overall trend in throughput values is in accordance with the pointer direction variability measure.

		Physical Demand		Mental Demand		Temporal Demand		Performance		Effort		Frustration	
Measure	N	df	χ^2	df	χ^2	df	χ^2	df	χ^2	df	χ^2	df	χ^2
Pointer Direction Variability	354	1	6.45 *	1	5.20 *	1	—	1	14.62 **	1	4.68 *	1	14.02 **
Hit Point Variability	354	1	—	1	4.27 *	1	—	1	12.93 **	1	—	1	13.43 **
Mean Alignment Distance	210	1	2.74 .	1	7.75 **	1	—	1	22.91 **	1	4.81 *	1	6.88 **

Table 2: Chi-square (χ^2) statistics for all significant ($p < .05$) and marginal ($p < .10$) effects of our measures on the NASA-TLX task load measures. Significance indicators: . $p < .10$, $^*p < .05$, $^{**}p < .01$.

		Eye			Head and Neck			Hand and Finger			Arm		
Measure	N	df	χ^2	η_p^2	df	χ^2	η_p^2	df	χ^2	η_p^2	df	χ^2	η_p^2
Gaze Direction Range	354	1	6.41 *	0.02	1	—	—	1	6.13 *	0.02	1	—	—
Head Direction Range	354	1	—	—	1	6.58 *	0.02	1	—	—	1	—	—
Palm Position Drift	210	1	—	—	1	3.72 .	0.02	1	5.42 *	0.03	1	4.51 *	0.02
Index Fingertip Variability	210	1	14.96 **	0.08	1	7.00 **	0.04	1	—	—	1	3.19 .	0.02

Table 3: Chi-square (χ^2) statistics and effect size (η_p^2) for all significant ($p < .05$) and marginal ($p < .10$) effects of our measures on the Borg CR-10 perceived exertion measures. Significance indicators: . $p < .10$, $^*p < .05$, $^{**}p < .01$.

		Ease of Use		Ease of Learning		Comfort		Sense of Agency	
Measure	N	df	χ^2	df	χ^2	df	χ^2	df	χ^2
Pointer Direction Variability	343	1	15.77 **	1	—	1	12.13 **	1	12.04 **
Hit Point Variability	343	1	12.76 **	1	—	1	8.84 **	1	13.60 **
Mean Alignment Distance	203	1	10.64 **	1	7.79 **	1	—	1	6.92 **

Table 4: Chi-square (χ^2) statistics for all significant ($p < .05$) effects of our measures on the additional user experience ratings. Significance indicators: $^*p < .05$, $^{**}p < .01$.

Using alignment-adjusted endpoints improves pointing accuracy. We computed the overall hit rate before and after using the alignment-adjusted endpoint for *Controller*, *Handray*, and *Ring*, the three techniques that did not use gaze as the pointer. The overall hit rate increased from 91.8% to 93.6% after using the alignment-adjusted endpoint to determine hit or miss. Among the 8170 trials, 196 trials changed from *miss* to *hit*, 52 trials changed from *hit* to *miss*, while 7922 trials remained unchanged. We found significant correlations between the *Endpoint Type* and the binary trial-level hit or miss ($\chi^2_{1,N=8170} = 19.96, p < .01$).

5.4 How Our Measures Characterize Pointing User Experience

In Section 5.2, we discussed overall trends of correlations between measure values and user ratings through the plot visualizations. In this section, we present how these measures characterize user experience by showing inferential analyses results on how they are correlated with the NASA-TLX task load measures, Borg CR-10 scale for perceived exertion, and additional user experience ratings of ease of use/learning, comfort, and sense of agency.

5.4.1 How Our Measures Reflect ***Task Load****.* We found significant correlations between our pointer measures and all six NASA-TLX task load measures except temporal demand. The correlations between our measures and the NASA-TLX task load measures are presented in Table 2. Specifically, lower pointer measure values were associated with higher task load—taking pointer direction variability and physical demand as an example (Figure 7), *Handray* had the lowest pointer direction variability ($M = 0.01, SD = 0.03$) (i.e., smallest amount of movement) yet the highest physical demand ($M = 3.5, SD = 1.8$) among the five techniques; in contrast, *Controller* had the highest pointer direction variability ($M = 0.04, SD = 0.02$) (i.e., largest amount of movement), yet the lowest physical demand ($M = 2.4, SD = 1.9$).

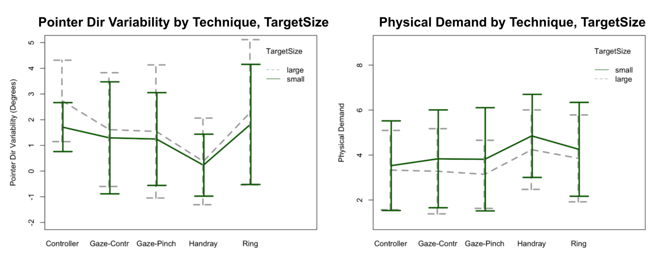

Figure 7: Interaction plots showing pointer direction variability and physical demand (error bar representing $M \pm SD$) by *Technique* and *Target Size*. The trends in the plot show that pointer direction variability and physical demand were generally negatively correlated.

5.4.2 How Our Measures Reflect ***Perceived Exertion****.* We found significant correlations between our measures and Borg CR-10 measures for perceived exertion. We present all significant ($p < .05$) and marginal ($p < .10$) correlations between our measures and perceived exertion in the eye, head and neck, hand and finger, and arm in Table 3. Specifically, lower measure values were associated with higher perceived exertion—taking gaze direction range and eye exertion as an example, *Gaze-Contr* and *Gaze-Pinch* had the lowest gaze direction range ($M = 0.02, SD = 0.04$ and $M = 0.02, SD = 0.03$) among the five techniques, and the highest eye exertion ($M = 3.1, SD = 2.3$ and $M = 3.1, SD = 2.2$) (Figure 8).

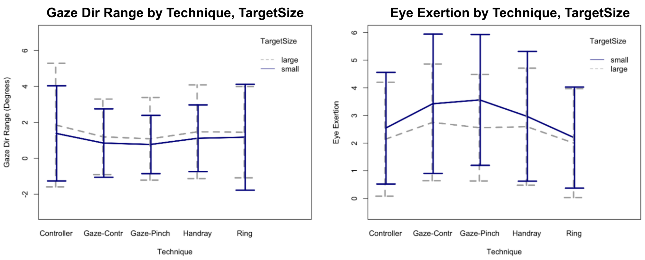

Figure 8: Interaction plots showing gaze direction range and perceived eye exertion (error bar representing $M \pm SD$) by *Technique* and *Target Size*. The trends in the plot show that gaze direction range and perceived eye exertion were generally negatively correlated.

5.4.3 How Our Measures Reflect ***Additional User Experience Ratings***. We found significant correlations between our measures and other user experience ratings including ease of use, ease of learning, comfort, and sense of agency. All significant ($p < .05$) correlations between our measures and the user experience ratings are presented in Table 4. Similar to task load measures, lower measure values (i.e., smaller amount of movements) were associated with degraded user experience—lower ease of use/learning, less comfort, and less sense of agency.

6 Discussion

6.1 Characterizing Performance and User Experience

In Section 5.4, we showed the correlations between our measures and existing pointing task measures. We demonstrated that our measures could effectively characterize pointing speed, accuracy, throughput, task load, perceived exertion, ease of use/learning, comfort, and sense of agency. Note that our measures do not require always-on sensor streaming, and only rely on a very small segment of sensor data between selection-start and selection-end—this makes our measures lightweight and suitable for runtime usage.

6.2 Findings

In this section, we discuss implications of our analysis results, and how we might understand pointing experiences through the lens of these measures.

Speed has a larger effect on overall performance than accuracy for the five techniques we evaluated. In Section 5.3, we found that techniques with lower measure values (i.e., more "stable" pointing) exhibited higher accuracy but lower throughput. This suggests that speed had a larger effect on overall performance than accuracy across the five techniques we evaluated. One possible explanation is the additional effort required for precise 3D pointing, which often involves holding the pointer steady—in contrast to mouse- or touch-based 2D input, which is easier to control.

Higher tolerance for less "stable" pointing improves the pointing user experience. In Section 5.3 and 5.4, we found that higher measure values (i.e., less "stable" pointing) were associated with lower task load, reduced fatigue, and better overall user experience. This highlights a key difference between 2D and 3D pointing: prior work on 2D pointing shows that less pointer movement corresponds to better fine motor control and performance—for touch-based systems, for example, a "perfect touch" would have zero drift [30, 31]. However, when it comes to 3D pointing, such "perfect" selections are much harder to achieve. Thus, unlike 2D pointing, our results suggest that greater tolerance for less "stable" pointing improves the user experience in 3D.

Lower latency and higher sampling rate might be more important than sensing accuracy for 3D pointing. It is important to note that the five techniques rely on different sensing systems, ranging from accelerometer and gyroscope to camera-based hand and gaze tracking.[5] While this may affect comparability, sensor characteristics are also integral to user experience. For example, participants described *Handray* as "laggy," requiring extra time to confirm target hover (P20), and reported feeling more in control when pointers "actually followed the movement" (P23). Our findings suggest that when trading off sensing accuracy and latency, techniques with slightly *lower accuracy* but *lower latency* may be preferable to those with *higher accuracy* but *higher latency*.

Different modalities don't work in isolation during 3D pointing. In Section 5.4.2, we found correlations not only within modalities (e.g., gaze direction range and eye exertion), but also across modalities. For example, gaze direction range was correlated with hand and finger exertion; mean alignment distance was correlated with task load. These results suggest that increased load in one modality may affect others. For instance, users may closely monitor hand ray-cast movements while controlling them, increasing eye exertion. Additionally, improved pointing accuracy with the alignment-adjusted endpoint measure (Section 5.3) further highlights the importance of coordination across modalities.

6.3 Design Implications

We include a summary of how pointing performance and user experience are associated with high and low measure values, i.e., the amount of movement during 3D pointing in Table 5. Overall, allowing for *more movements* during the "click" leads to better performance and user experience.

Measure Value	Performance			User Experience Ratings		
	Accuracy	Speed	Throughput	Task Load	Exertion	Additional
High	Low	High *	High *	Low *	Low *	High *
Low	High *	Low	Low	High	High	Low

Table 5: Pointing performance and user experience associated with high and low measure values, i.e., the amount of movement during 3D pointing. * indicates which option means "better" performance or user experience.

Our findings suggest that: (1) 3D pointing techniques should *improve tolerance for less "stable" pointing*; (2) 3D pointing techniques should *prioritize movement speed over pointing accuracy*; (3) sensing techniques for 3D pointing should *prioritize latency over sensing accuracy*; (4) 3D pointing interactions should *consider coordination among different modalities.*

[5] *Handray* operates at a lower sampling rate (60 Hz) than controller-, gaze-, and ring-based techniques (72 Hz), resulting in less captured movement and potentially contributing to observed differences in performance and user experience.

6.4 Limitations and Future Work

Our experiment evaluated only spherical targets of varying sizes on vertical task planes in front of the user. Future work could explore additional configurations, including target depth, shape, plane orientation, and larger angles from the center of view. While our measures could potentially be generalizable to other 3D pointing scenarios, e.g., augmented reality (AR) or mixed reality (MR), we validated them only in fully virtual environments; extending the evaluation to AR and MR is an important agenda for future work.

Our measure computation is dependent on sensor capabilities and detection algorithms. For example, hand movement capture relied on joint data from Meta OVRHand and OVRSkeleton, and our testbed used the default eye and hand tracking of the Meta Quest Pro. Tracking accuracy could be improved with specialized eye-tracking (e.g., Tobii Glasses) or motion-tracking systems (e.g., OptiTrack), which may influence targeting performance [15, 56]. Inherent differences in sensing techniques and tracking modalities may also limit the extent to which observed differences can be attributed to the input techniques themselves. Additionally, prior user experience with input modalities (e.g., controllers) may introduce bias, and future work could examine its relationship with task load.

Finally, we focused on pointing selection tasks in a controlled lab setting. Extending these measures to other interactions (e.g., object manipulation) and evaluating them in real-world contexts is an important direction, as practical tasks may involve concurrent use of modalities (e.g., gaze for browsing and selection), potentially increasing fatigue over time. For future work, field experiments with real-world tasks and validation with independent datasets are required to understand 3D pointing.

7 Conclusion

In this work, we presented eight pointing measures that characterize what happens *during* 3D pointing selection, covering pointer movements, user behavior in single modalities, and alignment between modalities. To validate these measures, we ran a study to collect pointing selection data with 36 participants and 5 different pointing techniques: *Controller*, *Gaze-Contr*, *Gaze-Pinch*, *Handray*, and *Ring*. Our analyses showed that our measures effectively characterized pointing speed, accuracy, throughput, task load, perceived exertion, ease of use/learning, comfort, and sense of agency. We found an association of less "stable" pointing with improved performance and user experience. We also found that using the alignment-adjusted measure improved the pointing performance. Our findings suggest that 3D pointing should prioritize speed over accuracy, improve tolerance for less stable pointing, and consider coordination between modalities.

Safe and Responsible Innovation Statement. The study was approved by the IRB and conducted in accordance with the organization's privacy and ethical guidelines. All study data were de-identified. The experiment was designed to minimize participant discomfort potentially caused by extended headset use.

References

[1] Ferran Argelaguet and Carlos Andujar. 2013. A survey of 3D object selection techniques for virtual environments. *Computers & Graphics* 37, 3 (2013), 121–136.

[2] Anil Ufuk Batmaz, Mohammad Rajabi Seraji, Johanna Kneifel, and Wolfgang Stuerzlinger. 2020. No jitter please: Effects of rotational and positional jitter on 3D mid-air interaction. In *Proceedings of the Future Technologies Conference*. Springer, 792–808.

[3] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2019. The effect of rotational jitter on 3D pointing tasks. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–6.

[4] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2019. Effects of 3D rotational jitter and selection methods on 3D pointing tasks. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 1687–1692.

[5] Xiaojun Bi, Yang Li, and Shumin Zhai. 2013. FFitts law: modeling finger touch with fitts' law. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. 1363–1372.

[6] Xiaojun Bi and Shumin Zhai. 2013. Bayesian touch: a statistical criterion of target selection with finger touch. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*. 51–60.

[7] Benjamin M Bolker, Mollie E Brooks, Connie J Clark, Shane W Geange, John R Poulsen, M Henry H Stevens, and Jada-Simone S White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in ecology & evolution* 24, 3 (2009), 127–135.

[8] Gunnar A Borg. 1982. Psychophysical bases of perceived exertion. *Medicine and science in sports and exercise* 14, 5 (1982), 377–381.

[9] Doug Bowman, Chadwick Wingrave, Joshua Campbell, and Vinh Ly. 2001. Using pinch gloves (tm) for both natural and abstract interaction techniques in virtual environments. (2001).

[10] Charlie S Burlingham, Naveen Sendhilnathan, Oleg Komogortsev, T Scott Murdison, and Michael J Proulx. 2024. Motor "laziness" constrains fixation selection in real-world tasks. *Proceedings of the National Academy of Sciences* 121, 12 (2024), e2302239121.

[11] Ishan Chatterjee, Robert Xiao, and Chris Harrison. 2015. Gaze+ gesture: Expressive, precise and targeted free-space interactions. In *Proceedings of the 2015 ACM on international conference on multimodal interaction*. 131–138.

[12] 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*. 772–773.

[13] Abigail Evans and Jacob Wobbrock. 2012. Taming wild behavior: The input observer for obtaining text entry and mouse pointing measures from everyday computer use. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. 1947–1956.

[14] Ajoy S Fernandes, T Scott Murdison, and Michael J Proulx. 2023. Leveling the playing field: A comparative reevaluation of unmodified eye tracking as an input and interaction modality for VR. *IEEE Transactions on Visualization and Computer Graphics* 29, 5 (2023), 2269–2279.

[15] Ajoy Savio Fernandes, T Scott Murdison, Immo Schuetz, Oleg Komogortsev, and Michael J Proulx. 2024. The Effect of Degraded Eye Tracking Accuracy on Interactions in VR. In *Proceedings of the 2024 Symposium on Eye Tracking Research and Applications*. 1–7.

[16] 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.

[17] Brigitte N Frederick. 1999. Fixed-, Random-, and Mixed-Effects ANOVA Models: A User-Friendly Guide for Increasing the Generalizability of ANOVA Results. (1999).

[18] 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*. 1–5.

[19] Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage publications Sage CA: Los Angeles, CA, 904–908.

[20] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[21] Donald Hedeker and Robert D Gibbons. 1994. A random-effects ordinal regression model for multilevel analysis. *Biometrics* (1994), 933–944.

[22] 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*. 1063–1072.

[23] Sture Holm. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* (1979), 65–70.

[24] Christian Holz and Patrick Baudisch. 2010. The generalized perceived input point model and how to double touch accuracy by extracting fingerprints. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. 581–590.

[25] Christian Holz and Patrick Baudisch. 2011. Understanding touch. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. 2501–2510.

[26] Faustina Hwang, Simeon Keates, Patrick Langdon, and John Clarkson. 2004. Mouse movements of motion-impaired users: A submovement analysis. In *Proceedings of the 5th International ACM SIGACCESS Conference on Computers and Accessibility*. 102–109.

[27] Sujin Jang, Wolfgang Stuerzlinger, Satyajit Ambike, and Karthik Ramani. 2017. Modeling cumulative arm fatigue in mid-air interaction based on perceived exertion and kinetics of arm motion. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 3328–3339.

[28] Simeon Keates, Faustina Hwang, Patrick Langdon, P John Clarkson, and Peter Robinson. 2002. Cursor measures for motion-impaired computer users. In *Proceedings of the 4th International ACM SIGACCESS Conference on Computers and Accessibility*. 135–142.

[29] Wolf Kienzle, Eric Whitmire, Chris Rittaler, and Hrvoje Benko. 2021. Electroring: Subtle pinch and touch detection with a ring. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–12.

[30] Junhan Kong, Mingyuan Zhong, James Fogarty, and Jacob O Wobbrock. 2021. New metrics for understanding touch by people with and without limited fine motor function. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–4.

[31] Junhan Kong, Mingyuan Zhong, James Fogarty, and Jacob O Wobbrock. 2022. Quantifying Touch: New Metrics for Characterizing What Happens During a Touch. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*. 1–13.

[32] 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*. 1–14.

[33] Ramon C Littell, PR Henry, and Clarence B Ammerman. 1998. Statistical analysis of repeated measures data using SAS procedures. *Journal of Animal Science* 76, 4 (1998), 1216–1231.

[34] Lu Lu, Pengshuai Duan, Xukun Shen, Shijin Zhang, Huiyan Feng, and Yong Flu. 2021. Gaze-pinch menu: Performing multiple interactions concurrently in mixed reality. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 536–537.

[35] Tiffany Luong, Yi Fei Cheng, Max Möbus, Andreas Fender, and Christian Holz. 2023. Controllers or Bare Hands? A Controlled Evaluation of Input Techniques on Interaction Performance and Exertion in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* (2023).

[36] Francisco Lopez Luro and Veronica Sundstedt. 2019. A comparative study of eye tracking and hand controller for aiming tasks in virtual reality. In *Proceedings of the 11th ACM Symposium on eye tracking research & applications*. 1–9.

[37] 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. *Proceedings of the ACM on Human-Computer Interaction* 6, ETRA (2022), 1–18.

[38] I Scott MacKenzie and Poika Isokoski. 2008. Fitts' throughput and the speed-accuracy tradeoff. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1633–1636.

[39] I Scott MacKenzie, Tatu Kauppinen, and Miika Silfverberg. 2001. Accuracy measures for evaluating computer pointing devices. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. 9–16.

[40] Päivi Majaranta and Kari-Jouko Räihä. 2002. Twenty years of eye typing: systems and design issues. In *Proceedings of the 2002 symposium on Eye tracking research & applications*. 15–22.

[41] 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*. 1–9.

[42] Martez E Mott, Radu-Daniel Vatavu, Shaun K Kane, and Jacob O Wobbrock. 2016. Smart touch: Improving touch accuracy for people with motor impairments with template matching. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. 1934–1946.

[43] 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*. 1–7.

[44] 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*. 99–108.

[45] Alexander Plopski, Teresa Hirzle, Nahal Norouzi, Long Qian, Gerd Bruder, and Tobias Langlotz. 2022. The eye in extended reality: A survey on gaze interaction and eye tracking in head-worn extended reality. *ACM Computing Surveys (CSUR)* 55, 3 (2022), 1–39.

[46] Ivan Poupyrev, Tadao Ichikawa, Suzanne Weghorst, and Mark Billinghurst. 1998. Egocentric object manipulation in virtual environments: empirical evaluation of interaction techniques. In *Computer graphics forum*, Vol. 17. Wiley Online Library, 41–52.

[47] Robin Schweigert, Valentin Schwind, and Sven Mayer. 2019. Eyepointing: A gaze-based selection technique. In *Proceedings of Mensch und Computer 2019*. 719–723.

[48] Valentin Schwind, Sven Mayer, Alexandre Comeau-Vermeersch, Robin Schweigert, and Niels Henze. 2018. Up to the finger tip: The effect of avatars on mid-air pointing accuracy in virtual reality. In *Proceedings of the 2018 annual symposium on computer-human interaction in play*. 477–488.

[49] Linda E Sibert and Robert JK Jacob. 2000. Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 281–288.

[50] 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*. 1161–1174.

[51] 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 (2004), 751–789.

[52] Marco Speicher, Anna Maria Feit, Pascal Ziegler, and Antonio Krüger. 2018. Selection-based text entry in virtual reality. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–13.

[53] Vildan Tanriverdi and Robert JK Jacob. 2000. Interacting with eye movements in virtual environments. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 265–272.

[54] Eduardo Velloso, Jayson Turner, Jason Alexander, Andreas Bulling, and Hans Gellersen. 2015. An empirical investigation of gaze selection in mid-air gestural 3D manipulation. In *Human-Computer Interaction–INTERACT 2015: 15th IFIP TC 13 International Conference, Bamberg, Germany, September 14-18, 2015, Proceedings, Part II 15*. Springer, 315–330.

[55] 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*. 1–15.

[56] Shu Wei, Desmond Bloemers, and Aitor Rovira. 2023. A preliminary study of the eye tracker in the meta quest pro. In *Proceedings of the 2023 ACM international conference on interactive media experiences*. 216–221.

[57] 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*. 1–10.

[58] Difeng Yu, Xueshi Lu, Rongkai Shi, Hai-Ning Liang, Tilman Dingler, Eduardo Velloso, and Jorge Goncalves. 2021. Gaze-supported 3d object manipulation in virtual reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.

[59] Shumin Zhai, Carlos Morimoto, and Steven Ihde. 1999. Manual and gaze input cascaded (MAGIC) pointing. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 246–253.

[60] Futian Zhang, Keiko Katsuragawa, and Edward Lank. 2022. Conductor: Intersection-based bimanual pointing in augmented and virtual reality. *Proceedings of the ACM on Human-Computer Interaction* 6, ISS (2022), 103–117.

[61] Yanxia Zhang, Sophie Stellmach, Abigail Sellen, and Andrew Blake. 2015. The costs and benefits of combining gaze and hand gestures for remote interaction. In *Human-Computer Interaction–INTERACT 2015: 15th IFIP TC 13 International Conference, Bamberg, Germany, September 14-18, 2015, Proceedings, Part III 15*. Springer, 570–577.

[62] Qian Zhou, George Fitzmaurice, and Fraser Anderson. 2022. In-depth mouse: Integrating desktop mouse into virtual reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.

A Full List of Measures We Explored

A.1 Target-Agnostic Measures

First, we formalize target-agnostic measures focusing on the movement traces themselves regardless of the target. These measures were inspired by target-agnostic metrics for characterizing touch behavior on 2D touchscreens [30, 31], aiming to cover modalities relevant to 3D pointing. For the hand-based measures, in particular, both overall movement such as palm positions and individual joint movements are considered to capture user behavior at different levels of granularity.

A.1.1 Active Pointer Measures. Let n be the total number of pointer events captured from selection-start to selection-end, inclusive. A pointing process can be defined as a sequence of pointer event $E_0, ..., E_{n-1}$, where selection-start is E_0 and selection-end is E_{n-1}. Then, for the i^{th} pointer event:

- $p_i = (p_{x,i}, p_{y,i}, p_{z,i})$ is the origin position of the ray-cast;
- $d_i = (d_{x,i}, d_{y,i}, d_{z,i})$ is the direction of the ray-cast;

- If available, (x_i, y_i, z_i) is the position of the pointer 3D location;
- t_i is the event's timestamp.

We define the angle change from d_i to d_j as $\Delta(d_i, d_j)$ and the Euclidean distance between 3D points p_i and p_j as $D(p_i, p_j)$.

Pointer Direction Drift. Pointer direction drift is the angle between the pointer ray-cast directions at selection-start and selection-end. Intuitively, this measure indicates how far away a pointer direction finished from where it started.

$$\text{Pointer Direction Drift} = \Delta(d_0, d_{n-1}) \in [0, \infty)$$

Pointer Direction Variability. Pointer direction variability is the cumulative change of pointing direction over the duration of selection. Intuitively, this measure indicates how "jittery" the pointing direction was during a selection.

$$\text{Pointer Direction Variability} = \sum_{i=1}^{n-1} \Delta(d_{i-1}, d_i) \in [0, \infty)$$

Pointer Direction Range. Pointer direction range is the angle between the two pointing directions that are farthest apart. Intuitively, this measure indicates the range of directions the pointing ray-cast covered throughout the selection.

$$\text{Pointer Direction Range} = \left(\max_{i,j \in 0,\dots,n-1} \Delta(d_i, d_j) \right) \in [0, \infty)$$

A.1.2 User Behavior Measures for Specific Modalities. We define gaze, head, hand and finger movement events captured from pointer down to pointer up as follows.

- Gaze Movement: $G_0, \dots, G_{n_g-1}$, where $G_i = (p_{g,i}, d_{g,i}) = (p_{x_g,i}, p_{y_g,i}, p_{z_g,i}, d_{x_g,i}, d_{y_g,i}, d_{z_g,i})$ is the 6 DoF gaze ray direction, and $p'_{g,i} = (p'x_g, i, p'y_g, i, p'z_g, i)$ is the 3D gaze point;
- Head Movement: $H_0, \dots, H_{n_h-1}$, where $H_i = (p_{h,i}, d_{h,i}) = (p_{x_h,i}, p_{y_h,i}, p_{z_h,i}, d_{x_h,i}, d_{y_h,i}, d_{z_h,i})$ is the 6 DoF head ray direction;
- Palm Position Movement: $P_0, \dots, P_{n_p-1}$, where $P_i = (p_{x_p,i}, p_{y_p,i}, p_{z_p,i})$ is the 3D palm position;
- Finger Joint Movement: $F_0, \dots, F_{n_f-1}$, where

$$F_i = \begin{bmatrix} f_{i,0} \\ f_{i,1} \\ \vdots \\ f_{i,m-1} \end{bmatrix} = \begin{bmatrix} p_{x_f,i,0} & p_{y_f,i,0} & p_{z_f,i,0} \\ p_{x_f,i,1} & p_{y_f,i,1} & p_{z_f,i,1} \\ \vdots & \vdots & \vdots \\ p_{x_f,i,m-1} & p_{y_f,i,m-1} & p_{z_f,i,m-1} \end{bmatrix}$$

and m is the number of joints in a hand skeleton captured. Specifically, $f_{i,\text{idx}} = (p_{x_f,i,\text{idx}}, p_{y_f,i,\text{idx}}, p_{z_f,i,\text{idx}})$ is the i-th position of the index fingertip during the movement.

Gaze Direction Drift. Gaze direction drift is the angle between the gaze ray-cast directions at selection-start and selection-end.

$$\text{Gaze Direction Drift} = \Delta(d_{g,0}, d_{g,n_g-1}) \in [0, \infty)$$

Gaze Direction Variability. Gaze direction variability is the cumulative change of gaze direction over the duration of selection.

$$\text{Gaze Direction Variability} = \sum_{i=1}^{n_g-1} \Delta(d_{g,i-1}, d_{g,i}) \in [0, \infty)$$

Gaze Direction Range. Gaze direction range is the angle between the two gaze directions that are farthest apart.

$$\text{Gaze Direction Range} = \left(\max_{i,j \in 0,\dots,n_g-1} \Delta(d_{g,i}, d_{g,j}) \right) \in [0, \infty)$$

Gaze Point Drift. Gaze point drift is the Euclidean distance between the gaze 3D points at selection-start and selection-end.

$$\text{Gaze Point Drift} = D(p'_{g,0}, p'_{g,n_g-1}) \in [0, \infty)$$

Gaze Point Variability. Gaze point variability is the total Euclidean distance of successive gaze point movements over the duration of selection.

$$\text{Gaze Point Variability} = \sum_{i=1}^{n_g-1} D(p'_{g,i-1}, p'_{g,i}) \in [0, \infty)$$

Gaze Point Range. Gaze point range is the Euclidean distance between the two gaze points that are farthest apart.

$$\text{Gaze Point Range} = \left(\max_{i,j \in 0,\dots,n_g-1} D(p'_{g,i}, p'_{g,j}) \right) \in [0, \infty)$$

Head Direction Drift. Head direction drift is the angle between the head directions at selection-start and selection-end.

$$\text{Head Direction Drift} = \Delta(d_{h,0}, d_{h,n_h-1}) \in [0, \infty)$$

Head Direction Variability. Head direction variability is the cumulative change of head direction over the duration of selection;

$$\text{Head Direction Variability} = \sum_{i=1}^{n_h-1} \Delta(d_{h,i-1}, d_{h,i}) \in [0, \infty)$$

Head Direction Range. Head direction range is the angle between the two head directions that are farthest apart.

$$\text{Head Direction Range} = \left(\max_{i,j \in 0,\dots,n_h-1} \Delta(d_{h,i}, d_{h,j}) \right) \in [0, \infty)$$

Palm Position Drift. Palm position drift is the Euclidean distance between the palm positions at selection-start and selection-end. Intuitively, this measure indicates how far away the user's hand moved from where it started.

$$\text{Palm Position Drift} = D(P_0, P_{n_p-1}) \in [0, \infty)$$

Palm Position Variability. Palm position variability is the total distance covered by successive palm movements from selection-start to selection-end. Intuitively, this measure indicates how "jittery" one's hand is during the selection.

$$\text{Palm Position Variability} = \sum_{i=1}^{n_p-1} D(P_{i-1}, P_i) \in [0, \infty)$$

Palm Position Range. Palm position range is the Euclidean distance between the most distant two palm positions. Intuitively, this measure indicates the spatial range of the palm position during selection.

$$\text{Palm Position Range} = \left(\max_{i,j \in 0,\dots,n_p-1} D(P_i, P_j) \right) \in [0, \infty)$$

Index Fingertip Drift. Index fingertip drift is the Euclidean distance between index fingertips at selection-start and selection-end;

$$\text{Index Fingertip Drift} = D(f_{0,\text{idx}}, f_{n_f-1,\text{idx}}) \in [0, \infty)$$

Index Fingertip Variability. Index fingertip variability is the total distance covered by successive index fingertip movements from selection-start to selection-end;

$$\text{Index Fingertip Variability} = \sum_{i=1}^{n_f-1} D(f_{i-1,\text{idx}}, f_{i,\text{idx}}) \in [0, \infty)$$

Index Fingertip Range. Index fingertip range is the Euclidean distance between the most distant two index fingertip positions.

$$\text{Index Fingertip Range} = \left(\max_{i,j \in 0,\ldots,n_f-1} D(f_{i,\text{idx}}, f_{j,\text{idx}}) \right) \in [0, \infty)$$

Average Hand Joint Drift. Average hand joint drift is the average Euclidean distance between each hand joint positions at selection-start and selection-end, across all hand joints.

$$\text{Average Hand Joint Drift} = \frac{1}{m} \sum_{k=0}^{m-1} D(f_{0,k}, f_{n_f-1,k}) \in [0, \infty)$$

Average Hand Joint Variability. Average hand joint variability is the average total distance covered by successive movements of each hand joint from selection-start to selection-end, across all hand joints.

$$\text{Average Hand Joint Variability} = \frac{1}{m} \sum_{k=0}^{m-1} \sum_{i=1}^{n_f-1} D(f_{i-1,k}, f_{i,k}) \in [0, \infty)$$

Average Hand Joint Range. Average hand joint range is the average Euclidean distance between the most distant two positions of each hand joint during the trace, across all hand joints.

$$\text{Average Hand Joint Range} = \frac{1}{m} \sum_{k=0}^{m-1} \left(\max_{i,j \in 0,\ldots,n_f-1} D(f_{i,k}, f_{j,k}) \right) \in [0, \infty)$$

A.2 Target-Aware Measures

With knowledge of the target size and position, our measures also look into whether the pointer ray hits the target and the movement traces of the hit points with both the target object and the task plane. The target-aware measures were inspired by MacKenzie et al.'s accuracy measures characterizing what happens during mouse pointing on 2D graphical UIs [39], and extended to characterize pointer interactions with 3D targets and 2D task planes, as well as alignment between multiple modalities.

A.2.1 Task Plane Definition. When a user is interacting with menu items or 2D UI components on a flat plane in the 3D space, the task plane is the 2D plane containing the UI components; when the target becomes a 3D object and there is no explicit 2D task plane for the pointing task, our measures consider both the flat surface the user is facing where the target center resides and the cylindrical surface around the user (i.e. around the head-corrected world coordinate y-axis) as potential task planes.

A.2.2 Target Hit Measures. For the i^{th} pointer event, I_i is 1 when the pointer ray-cast hits the target object and 0 when it doesn't. Similarly, for the i^{th} gaze event (when gaze is not used as the active pointer), I'_i is 1 when the gaze ray hits the target object and 0 when it doesn't.

Target Hit Ratio. Target hit ratio is the percentage of pointer events that hit the target.

$$\text{Target Hit Ratio} = \frac{\sum_{i=0}^{n-1} I_i}{n} \times 100\%$$

Target Focus Ratio. Target focus ratio is the percentage of gaze rays that hit the target.

$$\text{Target Focus Ratio} = \frac{\sum_{i=0}^{n-1} I'_i}{n} \times 100\%$$

Target Re-Entry. Target re-entry is defined as the number of times the active pointer enters the target.

Focus Re-Entry. Focus re-entry is defined as the number of times the gaze ray enters the target.

A.2.3 Task Plane Hit Measures. For the i^{th} pointer event, $\text{Hit}_i = (x_{\text{hit},i}, y_{\text{hit},i})$ is the position where the pointer ray-cast hits the task plane, in the task plane 2D coordinates. Similarly, for the i^{th} gaze event (when gaze is not used as the active pointer), $\text{Hit}'_i = (x'_{\text{hit},i}, y'_{\text{hit},i})$ is the position where the gaze ray hits the task plane, in the task plane 2D coordinates. We define the hit point measures characterizing pointer movement on the 2D task plane as below.

Hit Point Drift. Hit point drift is the Euclidean distance between the pointer ray-cast hit point with the task plane at selection-start and selection-end, in the task plane 2D coordinates.

$$\text{Hit Point Drift} = D(\text{Hit}_0, \text{Hit}_{n-1}) \in [0, \infty)$$

Hit Point Variability. Hit point variability is the total distance covered by successive hit point movements from selection-start to selection-end, in the task plane 2D coordinates.

$$\text{Hit Point Variability} = \sum_{i=1}^{n-1} D(\text{Hit}_{i-1}, \text{Hit}_i) \in [0, \infty)$$

Hit Point Range. Hit point range is the Euclidean distance between the two most distant hit points, in the task plane 2D coordinates.

$$\text{Hit Point Range} = \left(\max_{i,j \in 0,\ldots,n-1} D(\text{Hit}_i, \text{Hit}_j) \right) \in [0, \infty)$$

A.2.4 Alignment Measures. The target-aware measures also look into how well different modalities coordinate in a pointing gesture, for example, the alignment between the controller or hand ray-cast and the gaze ray.

Mean Alignment Distance. Alignment distance is the average distance between the active pointer hit point and the gaze hit point on the task plane.

$$\text{Mean Alignment Distance} = \frac{1}{n} \sum_{i=0}^{n-1} D(\text{Hit}_i, \text{Hit}'_i) \in [0, \infty)$$

Alignment Distance Deviation. Alignment distance deviation is the standard deviation of the distance between the active pointer hit point and the gaze hit point on the task plane, where μ_{AD} is the mean alignment distance.

$$\text{Alignment Distance Deviation} = \sqrt{\frac{1}{n}\sum_{i=0}^{n-1}[D(\text{Hit}_i, \text{Hit}'_i) - \mu_{\text{AD}}]^2} \in [0, \infty)$$

Alignment-Adjusted Endpoint. Alignment-adjusted endpoint is the position of the 3D pointer location where the active pointer and gaze hit point locations are closest on the task plane.

$$\text{Alignment-Adjusted Endpoint} =$$

$$(x_{i^*}, y_{i^*}, z_{i^*}) \text{ where } i^* = \underset{i=0,\ldots,n-1}{\operatorname{argmin}} D(\text{Hit}_i, \text{Hit}'_i)$$

A.3 Measure Collection and Evaluation

For each trial, we collected the active pointer directions, gaze positions and directions, head positions and directions, hand and finger joint positions, and target size and position, using the testbed (Section 4.3). The 31 measures were then calculated from the collected trials, and analyzed using methods described in Section 4.6.4. Note that while more than 8 measures exhibited significant effects as shown in Appendix E, we selected a representative subset by removing measures that captured similar aspects of the same modality and showed similar significance patterns, in order to maintain interpretability and avoid repetition.

B Interaction Plots for Task Load, Perceived Exertion, and Additional User Ratings

In Figures 9-11, we plot the mean and standard deviation of each rating question (NASA-TLX, Borg CR-10, and other user experience questions) by *Technique* and *Target Size*.

C Interaction Plots for All Measures

In Figures 12-18, we plot the mean and standard deviation of each measure by *Technique* and *Target Size*.

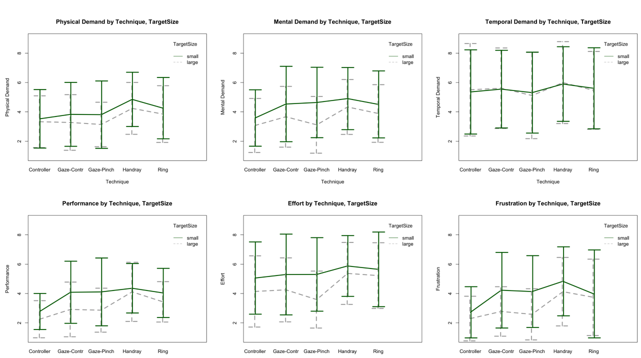

Figure 9: Interaction plots showing NASA-TLX task load ratings (error bar representing $M \pm SD$) of *Technique* × *Target Size*.

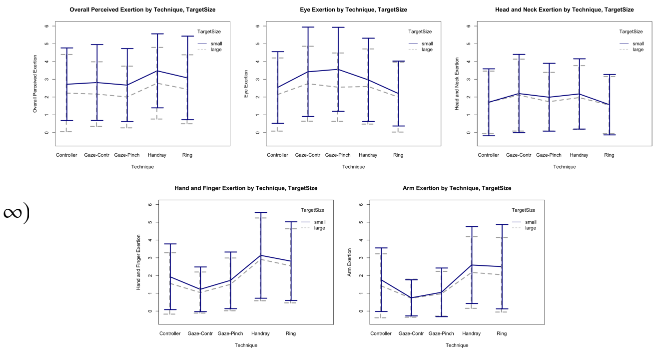

Figure 10: Interaction plots showing Borg CR-10 perceived exertion ratings (error bar representing $M \pm SD$) of *Technique* × *Target Size*.

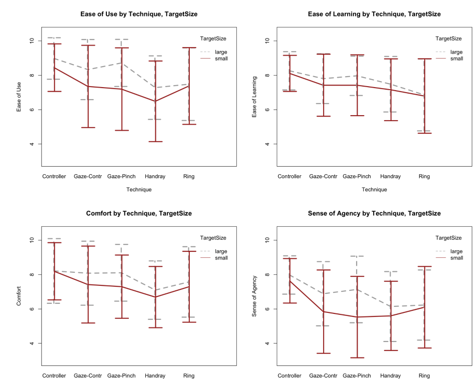

Figure 11: Interaction plots showing additional user experience ratings (error bar representing $M \pm SD$) of *Technique* × *Target Size*.

D Full Omnibus Test Results Comparing the Measures Across Techniques

We present the χ^2-statistics of the omnibus tests for our selected measures in Table 7. Alignment-adjusted endpoint is omitted from the comparison because it is a 3D vector rather than a scalar measure. We omit the full pairwise comparison results here due to the large number of combinations.

E Full Analysis Results of All 31 Measures

Table 6 shows all significant effects of *Technique*, *Target Size*, and their interaction on each measure. Table 8 shows correlations between each NASA-TLX task load measure and each measure. Table 9 shows correlations between Borg CR-10 perceived exertion scores and each measure. Table 10 shows correlations between additional user experience ratings and each measure. As the initial analyses

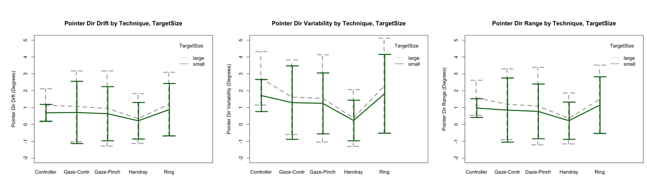

Figure 12: Interaction plots showing active pointer measures (error bar representing $M \pm SD$) of *Technique* × *Target Size*.

of the hit point measures yielded the same results for the flat and cylindrical task planes, we conducted the remaining analyses using only the flat task plane.

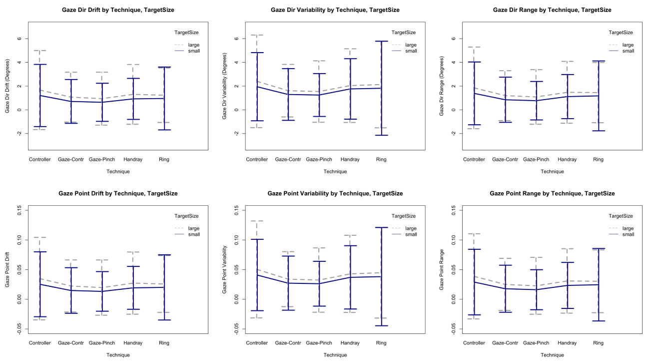

Figure 13: Interaction plots showing gaze measures (error bar representing $M \pm SD$) of *Technique* × *Target Size*.

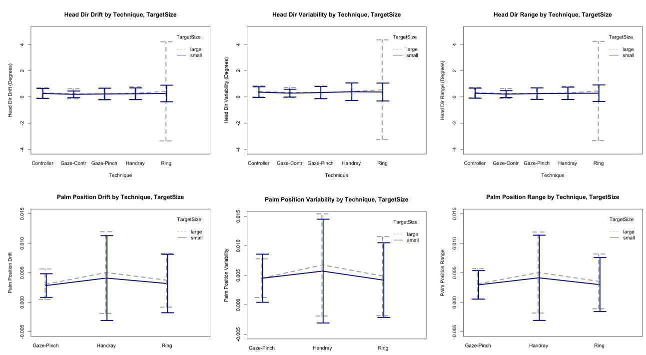

Figure 14: Interaction plots showing head and palm position measures (error bar representing $M \pm SD$) of *Technique* × *Target Size*.

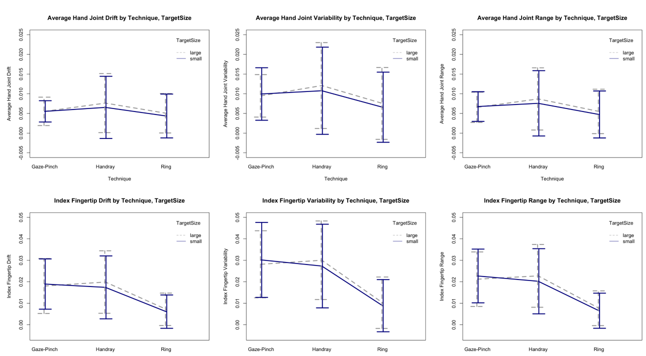

Figure 15: Interaction plots showing finger joint measures (error bar representing $M \pm SD$) of *Technique* × *Target Size*.

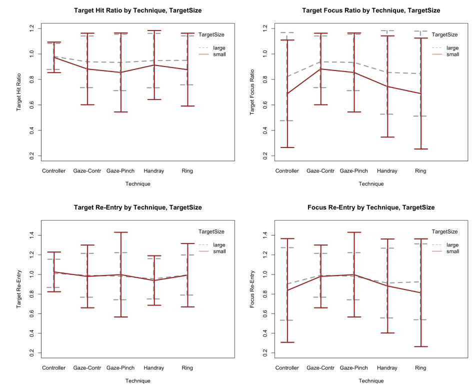

Figure 16: Interaction plots showing target hit and re-entry measures (error bar representing $M \pm SD$) of *Technique* × *Target Size*.

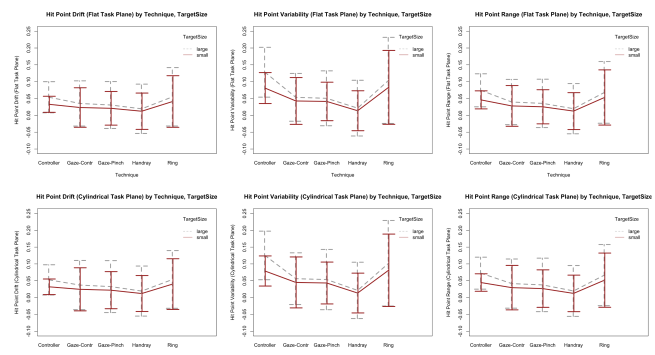

Figure 17: Interaction plots showing task plane hit point measures (error bar representing $M \pm SD$) of *Technique* × *Target Size*.

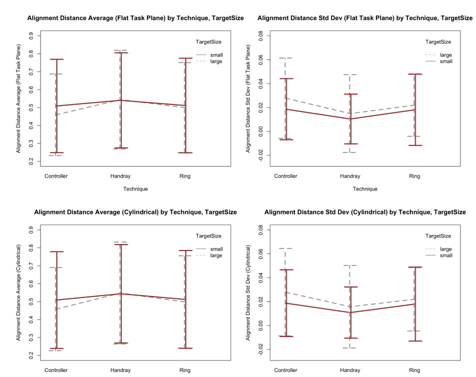

Figure 18: Interaction plots showing eye and pointer alignment measures (error bar representing $M \pm SD$) of *Technique* × *Target Size*.

		Technique			***Target Size***			***Technique*** × ***Target Size***		
Measure	N	df	χ^2	η_p^2	df	χ^2	η_p^2	df	χ^2	η_p^2
Active Pointer Measures										
Pointer Direction Drift	13885	4	6075.30 **	0.310	1	383.95 **	0.030	4	36.12 **	0.003
Pointer Direction Variability	13885	4	20674.40 **	0.600	1	378.37 **	0.030	4	112.51 **	0.008
Pointer Direction Range	13885	4	12710.90 **	0.480	1	451.32 **	0.030	4	74.74 **	0.005
Eye Movement Measures										
Gaze Direction Drift	13885	4	234.46 **	0.020	1	119.77 **	0.008	4	—	—
Gaze Point Drift	13885	4	242.34 **	0.020	1	125.81 **	0.009	4	—	—
Gaze Direction Variability	13885	4	237.49 **	0.020	1	36.06 **	0.002	4	10.01 *	0.001
Gaze Point Variability	13885	4	247.12 **	0.020	1	40.25 **	0.003	4	—	—
Gaze Direction Range	13885	4	265.03 **	0.020	1	101.13 **	0.007	4	—	—
Gaze Point Range	13885	4	273.14 **	0.020	1	106.56 **	0.008	4	—	—
Head Movement Measures										
Head Direction Drift	13885	4	318.32 **	0.020	1	64.98 **	0.005	4	12.89 *	0.001
Head Direction Variability	13885	4	281.59 **	0.020	1	47.78 **	0.003	4	15.77 **	0.001
Head Direction Range	13885	4	323.09 **	0.020	1	74.21 **	0.005	4	14.22 **	0.001
Hand Movement Measures										
Palm Position Drift	8095	2	314.59 **	0.040	1	—	—	2	15.87 **	0.002
Palm Position Variability	8095	2	489.32 **	0.060	1	—	—	2	28.90 **	0.004
Palm Position Range	8095	2	428.47 **	0.050	1	—	—	2	23.50 **	0.003
Avg Hand Joint Drift	8095	2	889.99 **	0.100	1	—	—	2	63.59 **	0.008
Avg Hand Joint Variability	8095	2	1835.53 **	0.190	1	4.69 *	0.005	2	54.18 **	0.007
Avg Hand Joint Range	8095	2	1470.84 **	0.150	1	4.94 *	0.008	2	70.73 **	0.009
Index Fingertip Drift	8095	2	4466.30 **	0.360	1	14.50 **	0.003	2	69.71 **	0.009
Index Fingertip Variability	8095	2	9412.21 **	0.540	1	12.72 **	0.004	2	73.93 **	0.009
Index Fingertip Range	8095	2	8024.82 **	0.500	1	18.19 **	0.004	2	92.41 **	0.010
Task Plane Hit Point Measures										
Hit Point Drift (Flat)	13885	4	4700.39 **	0.250	1	393.47 **	0.030	4	36.43 **	0.003
Hit Point Drift (Cylindrical)	13885	4	4489.84 **	0.240	1	390.90 **	0.030	4	36.86 **	0.003
Hit Point Variability (Flat)	13885	4	15718.37 **	0.530	1	382.87 **	0.030	4	111.18 **	0.008
Hit Point Variability (Cylindrical)	13885	4	15446.31 **	0.530	1	382.23 **	0.030	4	113.09 **	0.008
Hit Point Range (Flat)	13885	4	9868.16 **	0.410	1	457.25 **	0.030	4	73.43 **	0.005
Hit Point Range (Cylindrical)	13885	4	9549.05 **	0.410	1	456.75 **	0.030	4	75.23 **	0.005
Alignment Measures										
Mean Alignment Distance	8170	2	82.40 **	0.010	1	31.43 **	0.001	2	20.53 **	0.003
Alignment Distance Deviation	8170	2	206.99 **	0.020	1	91.54 **	0.010	2	21.77 **	0.003

Table 6: Chi-square (χ^2) statistics of all significant ($p < .05$) effects of Technique, Target Size, and their interaction on each measure. Significance indicators: $^*p < .05$, $^{**}p < .01$.

		Technique			***Target Size***			***Technique* × *Target Size***		
Measure	N	df	χ^2	η_p^2	df	χ^2	η_p^2	df	χ^2	η_p^2
Gaze Direction Range	13885	4	265.03 **	0.020	1	101.13 **	0.007	4	—	—
Head Direction Range	13885	4	323.09 **	0.020	1	74.21 **	0.005	4	14.22 **	0.001
Palm Position Drift	8095	2	314.59 **	0.040	1	—	—	2	15.87 **	0.002
Index Fingertip Variability	8095	2	9412.21 **	0.540	1	12.72 **	0.004	2	73.93 **	0.009
Pointer Direction Variability	13885	4	20674.40 **	0.600	1	378.37 **	0.030	4	112.51 **	0.008
Hit Point Variability	13885	4	15718.37 **	0.530	1	382.87 **	0.030	4	111.18 **	0.008
Mean Alignment Distance	8170	2	82.40 **	0.010	1	31.43 **	0.001	2	20.53 **	0.003

Table 7: Chi-square (χ^2) statistics of all significant ($p < .05$) effects of Technique, Target Size, and their interaction on each measure. Significance indicators: $^*p < .05$, $^{**}p < .01$.

		Physical Demand		**Mental Demand**		**Temporal Demand**		**Performance**		**Effort**		**Frustration**	
Measure	N	df	χ^2	df	χ^2	df	χ^2	df	χ^2	df	χ^2	df	χ^2
Active Pointer Measures													
Pointer Direction Drift	354	1	4.39 *	1	3.76 .	1	—	1	11.93 **	1	4.69 *	1	9.69 **
Pointer Direction Variability	354	1	6.45 *	1	5.20 *	1	—	1	14.62 **	1	4.68 *	1	14.02 **
Pointer Direction Range	354	1	4.85 *	1	4.34 *	1	—	1	13.35 **	1	4.36 *	1	11.71 **
Task Plane Hit Point Measures													
Hit Point Drift	354	1	—	1	2.77 .	1	—	1	9.54 **	1	—	1	8.38 **
Hit Point Variability	354	1	—	1	4.27 *	1	—	1	12.93 **	1	—	1	13.43 **
Hit Point Range	354	1	—	1	3.35 .	1	—	1	11.21 **	1	—	1	10.67 **
Alignment Measures													
Mean Alignment Distance	210	1	2.74 .	1	7.75 **	1	—	1	22.91 **	1	4.81 *	1	6.88 **
Alignment Distance Deviation	210	1	—	1	5.86 *	1	—	1	5.96 *	1	5.63 *	1	9.12 **

Table 8: Chi-square (χ^2) statistics for all significant ($p < .05$) and marginal ($p < .10$) effects of our measures on the NASA-TLX task load measures. Significance indicators: $.p < .10$, $^*p < .05$, $^{**}p < .01$.

		Eye			Head and Neck			Hand and Finger			Arm		
Measure	N	df	χ^2	η_p^2	df	χ^2	η_p^2	df	χ^2	η_p^2	df	χ^2	η_p^2
Eye Movement Measures													
Gaze Direction Drift	354	1	8.14 **	0.020	1	—	—	1	5.61 *	0.020	1	—	—
Gaze Direction Variability	354	1	3.17 .	0.009	1	—	—	1	7.63 **	0.020	1	—	—
Gaze Direction Range	354	1	6.41 *	0.020	1	—	—	1	6.13 *	0.020	1	—	—
Head Movement Measures													
Head Direction Drift	354	1	—	—	1	6.82 **	0.020	1	—	—	1	—	—
Head Direction Variability	354	1	—	—	1	5.57 *	0.020	1	—	—	1	—	—
Head Direction Range	354	1	—	—	1	6.58 *	0.020	1	—	—	1	—	—
Hand Movement Measures													
Palm Position Drift	210	1	—	—	1	3.72 .	0.020	1	5.42 *	0.030	1	4.51 *	0.020
Palm Position Variability	210	1	—	—	1	6.99 **	0.040	1	7.26 **	0.040	1	3.04 .	0.010
Palm Position Range	210	1	—	—	1	5.91 *	0.030	1	6.04 *	0.030	1	3.17 .	0.020
Avg Hand Joint Drift	210	1	3.60 .	0.020	1	5.98 *	0.030	1	2.86 .	0.010	1	—	—
Avg Hand Joint Variability	210	1	9.23 **	0.050	1	8.19 **	0.040	1	—	—	1	—	—
Avg Hand Joint Range	210	1	6.69 **	0.030	1	7.11 **	0.040	1	—	—	1	—	—
Index Fingertip Drift	210	1	10.78 **	0.060	1	3.72 .	0.020	1	—	—	1	4.11 *	0.020
Index Fingertip Variability	210	1	14.96 **	0.080	1	7.00 **	0.040	1	—	—	1	3.19 .	0.020
Index Fingertip Range	210	1	13.01 **	0.070	1	5.16 *	0.030	1	—	—	1	4.67 *	0.020
Alignment Measures													
Mean Alignment Distance	210	1	9.85 **	0.050	1	—	—	1	34.00 **	0.160	1	26.58 **	0.130
Alignment Distance Deviation	210	1	10.95 **	0.060	1	—	—	1	48.92 **	0.220	1	34.94 **	0.190

Table 9: Chi-square (χ^2) statistics for all significant ($p < .05$) and marginal ($p < .10$) effects of our measures on the Borg CR-10 perceived exertion measures. Significance indicators: . $p < .10$, $^*p < .05$, $^{**}p < .01$.

		Ease of Use		Ease of Learning		Comfort		Sense of Agency	
Measure	N	df	χ^2	df	χ^2	df	χ^2	df	χ^2
Active Pointer Measures									
Pointer Direction Drift	343	1	12.62 **	1	—	1	9.26 **	1	9.04 **
Pointer Direction Variability	343	1	15.77 **	1	—	1	12.13 **	1	12.04 **
Pointer Direction Range	343	1	14.12 **	1	—	1	10.63 **	1	10.71 **
Task Plane Hit Point Measures									
Hit Point Drift	343	1	8.77 **	1	—	1	5.29 *	1	10.02 **
Hit Point Variability	343	1	12.76 **	1	—	1	8.84 **	1	13.60 **
Hit Point Range	343	1	10.54 **	1	—	1	6.79 **	1	12.00 **
Alignment Measures									
Mean Alignment Distance	203	1	10.64 **	1	7.79 **	1	—	1	6.92 **
Alignment Distance Deviation	203	1	19.04 **	1	—	1	4.06 *	1	15.88 **

Table 10: Chi-square (χ^2) statistics for all significant ($p < .05$) effects of our measures on the additional user experience ratings. Significance indicators: $^*p < .05$, $^{**}p < .01$.