Exploring and Modeling the Effects of Eye-Tracking Accuracy and Precision on Gaze-Based Steering in Virtual Environments

Xuning Hu, Yichuan Zhang, Yushi Wei, Liangyuting Zhang, Yue Li, Wolfgang Stuerzlinger and Hai-Ning Liang Member, IEEE

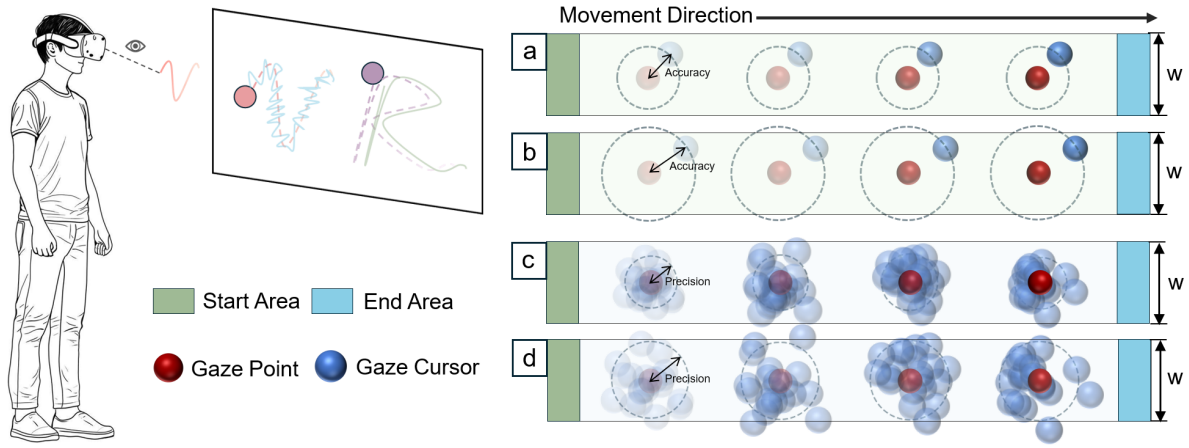

Fig. 1: The left side demonstrates our experimental setup, where participants complete a steering task using a gaze-based cursor in a VR environment. (a–b) illustrate simulated spatial inaccuracy by introducing a fixed offset between the user's actual gaze point and the displayed gaze cursor, representing systematic calibration errors (see Sec. 2.3 for formulas and details). (c–d) depict spatial imprecision, modeled by adding Gaussian noise to the gaze cursor position to reflect reduced stability in gaze estimation (see Sec. 2.3 for formulas and details).

Abstract—Recent advances in eye-tracking technology have positioned gaze as an efficient and intuitive input method for Virtual Reality (VR), offering a natural and immersive user experience. As a result, gaze input is now leveraged for fundamental interaction tasks such as selection, manipulation, crossing, and steering. Although several studies have modeled user steering performance across various path characteristics and input methods, our understanding of gaze-based steering in VR remains limited. This gap persists because the unique qualities of eye movements—involving rapid, continuous motions—and the variability in eye-tracking make findings from other input modalities nontransferable to a gaze-based context, underscoring the need for a dedicated investigation into gaze-based steering behaviors and performance. To bridge this gap, we present two user studies to explore and model gaze-based steering. In the first one, user behavior data are collected across various path characteristics and eye-tracking conditions. Based on this data, we propose four refined models that extend the classic Steering Law to predict users' movement time in gaze-based steering tasks, explicitly incorporating the impact of tracking quality. The best-performing model achieves an adjusted R^2 of 0.956, corresponding to a 16% improvement in movement time prediction. This model also yields a substantial reduction in AIC (from 1550 to 1132) and BIC (from 1555 to 1142), highlighting improved model quality and better balance between goodness of fit and model complexity. Finally, data from a second study with varied settings, such as a different eye-tracking sampling rate, illustrate the strong robustness and predictability of our models. Finally, we present scenarios and applications that demonstrate how our models can be used to design enhanced gaze-based interactions in VR systems.

Index Terms—Human Performance Modeling, Virtual Reality, Eye-Tracking Data Quality, Gaze-Based Interaction, Steering Law.

1 Introduction

A well-designed virtual reality (VR) system typically integrates multiple interaction methods—such as selection, dragging, crossing, and steering—to provide an efficient, seamless user experience. Among these methods, steering, i.e., guiding an object along a predefined path from a starting point to an endpoint, has been widely applied in areas such as gaming [20], menu navigation [19], and simulated driving [57]. The Steering Law [1] models how path characteristics influence user performance in steering tasks and helps understand user steering behaviors more deeply. Previous studies have confirmed this law's applicability to a wide range of hand-based input methods, including stylus- [22], bare-hand- [52], and controller-based interactions [16].

Although the Steering Law has been validated for various hand-based input methods, recent advances in eye-tracking technology now afford the possibility of performing steering tasks through gaze-based input. Thus, such gaze-based interaction has been given increased

- *Xuning Hu and Yichuan Zhang are with Xi'an Jiaotong-Liverpool University and the Hong Kong University of Science and Technology (Guangzhou).*
- *Yushi Wei is with the Hong Kong University of Science and Technology (Guangzhou).*
- *Liangyuting Zhang and Yue Li are with Xi'an Jiaotong-Liverpool University.*
- *Wolfgang Stuerzlinger is with Simon Fraser University.*
- *Hai-Ning Liang is with the Hong Kong University of Science and Technology (Guangzhou). E-mail: hainingliang@hkust-gz.edu.cn (corresponding author).*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

attention by the human-computer interaction community. Advances in eye-tracking technology now enable VR headsets to incorporate lightweight tracking systems, allowing users to interact with virtual environments (VEs) solely through their gaze. Prior research suggests that gaze-based input can be both intuitive and immersive while maintaining high efficiency [7, 28, 44]. Moreover, gaze-based methods offer genuine hands-free operation [24, 25], which is particularly beneficial in confined environments or situations where the users' hands are otherwise occupied.

Although numerous studies have examined steering tasks and gaze-based interaction separately, a comprehensive understanding of gaze-based steering in virtual environments remains limited. This gap exists because findings from other input methods—including bare-hand, controller, and head-tracking—cannot simply be generalized to gaze-based steering. Two key factors contribute to this. First, eye movements differ markedly from other input modalities: whereas hand or head movements tend to be relatively stable, the eyes exhibit constant motion—even during fixation—due to tremors, drifts, and microsaccades [10]. This inherent instability leads to performance outcomes and user experiences distinct from those observed with other input methods [39]. Second, analyzing gaze-based steering requires careful consideration of tracking systems and the quality of the data captured. Factors such as eye-tracking algorithms [26], individual pupil characteristics [36], target location [11], and scene brightness [51] can all influence tracking fidelity. Prior work further shows that fluctuations in tracking quality significantly affect both user performance and overall experience [12].

To address this gap, we conducted two user studies aimed at understanding gaze-based steering behavior and extending the classic Steering Law to enable accurate prediction of users' movement time in steering tasks while explicitly accounting for tracking-quality variations. In the first study (N = 24), participants performed steering tasks under varying conditions, e.g., differences in tracking quality and path characteristics, using the Meta Quest Pro, which features a proprietary eye-tracking system. We recorded movement time, success rate, and average movement speed *(see Sec. 4.7)*. Based on the data, we tested the original Steering Law for gaze-based tasks and quantified the impact of tracking quality, ultimately yielding four enhanced predictive models *(see Sec. 5.4)*. Subsequently, we conducted a second study (N = 14) using the Pico headset equipped with Tobii eye tracking to further validate these models by collecting additional behavioral data across different tracking quality conditions *(see Sec. 6)*. Our findings empirically demonstrate that, compared to previously established models, these refined models deliver superior predictive accuracy for both movement time and average movement speed.

In summary, the primary contributions of our work include:

- An empirical verification of the Steering Law's applicability to gaze-based steering tasks in VR.
- A systematic investigation and quantification of the effects of eye-tracking quality in VR systems.
- The development and validation of four novel predictive models that significantly enhance model performance in predicting movement time, with 16% higher adjusted R^2 and AIC and BIC reduced from 1550 to 1132 and from 1555 to 1142, accounting explicitly for gaze characteristics.
- A robustness evaluation of these predictive models across various eye-tracking qualities and sampling rates, providing practical guidelines for gaze-based interaction design.

2 Related Work

This section reviews related work in three areas. First, we discuss foundational HCI models, including Fitts' Law and the Steering Law. Second, we examine gaze-based techniques, highlighting interaction patterns and behavioral characteristics. Finally, we review work on eye-tracking quality, which is characterized by accuracy and precision. In particular, we focus on the factors that affect tracking performance, the experimental implications of these variations, and the methods used to simulate tracking errors.

2.1 Modeling Human Behavior in Pointing and Steering

To quantify and predict human performance in pointing tasks, Fitts' Law is a fundamental and widely used probabilistic model based on human behavioral patterns [27]. The movement time (MT) in Fitts' law refers to the time required to point at and select a target, which is influenced by the index of difficulty (ID). The ID is determined by two basic task parameters: target width (W) and target amplitude (A). The equation below expresses the relationship between MT and the task parameters:

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

W represents the width of the target, and A denotes the amplitude, i.e., the distance between the starting position and the target. The coefficients a and b are empirically determined through regression analysis.

Inspired by Fitts' law, the Steering Law focuses on interpreting behavioral patterns and performance in steering tasks, which involve users navigating through paths constrained by limited width and length [2,57]. Behavior and performance are typically represented by the movement time (MT) metric, which refers to the time required to steer from the start to the end point [1]. The relationship between MT, width (W), and length (A) can be expressed as [1]:

$$MT = a + b \cdot \int_C \frac{ds}{W(s)} \tag{2}$$

where a and b represent empirical constants, s denotes a specific position along path C, and $W(s)$ refers to the width at the position s. Assuming that path c maintains a constant width, Eq. (2) can be further simplified to:

$$MT = a + b \cdot \frac{A}{W} \tag{3}$$

Thus, the ds and $W(s)$ can be replaced by the independent value of path length A and path constant width W.

In real-world applications, steering behavior is often affected by additional system-level factors beyond path geometry. Relevant to our work, previous research has developed various extended models and explored how some unavoidable factors impact user performance by integrating them into predictive models. For example, Wei et al. [50] conducted experiments with eight commonly encountered device frame rates in VR to quantify the effect of frame rate on path steering tasks. Their findings demonstrate that the impact of frame rate on movement time follows an exponential trend. Consequently, they proposed a model in which movement time is influenced not only by the amplitude-to-width ratio of the path but also by an exponential function of the frame rate. To address unresolved issues associated with device delays, Yamanaka and Stuerzlinger [56] further investigated the effect of latency on path steering performance. They defined latency as a combination of inherent device delays and delays introduced by human factors. Their findings indicate that movement time increases proportionally with latency and that latency interacts with the amplitude and width of the path, suggesting a compound effect on user performance.

More relevant to our work, previous research has primarily focused on examining the performance of gaze-based input in selection tasks and how to model it using Fitts' Law. However, the predictive power of Fitts' Law for gaze input has been less clear. Zhai [60] found a relatively low fit for gaze, with a value of 0.75, while Vertegaal et al. [47] reported a higher fit of 0.86 for eye clicks. Miniotas et al. [32] reported the highest fit of 0.98. Based on the characteristics of eye movements, Zhang et al. [62] identified two main factors contributing to the instability of model fitting. The first factor is attributed to the inevitable micro-saccadic movements of the eyes, which prevent the full utilization of the target width. To address this, they proposed to subtract an empirical constant $\mu = 11.2$ from the original target width. The second factor stems from prior research suggesting that the contribution of target distance A to selection time is relatively low. Therefore, they introduced an empirical term $e^{\lambda A}$, with $\lambda = 0.00052$, to reduce the

impact of A on selection time. Based on this, they proposed an adapted version of Fitts' Law as follows:

$$MT = a + b \cdot \frac{e^{\lambda A}}{W - \mu} \tag{4}$$

Additionally, one recent study has applied the steering law to gaze-based performance in virtual environments [17]. While showing good predictability, the work did not take into account tracking quality—a known factor affecting user performance and experience (see Sec. 2.3).

2.2 Gaze-Based Interaction in VR

With the increasing integration of eye-tracking technology, many VR head-mounted displays (HMDs), such as the Apple Vision Pro, Quest Pro, and HTC Vive, now incorporate eye-tracking as a standard input method. Eye-tracking has been extensively explored and applied in various scenarios as a standalone input method or in combination with other input modalities, including in games [48], text input [24], and selection [30]. Scenarios most pertinent to our work include gaze-based target navigation [21], target manipulation [58], and menu steering interactions [19].

Oyckoya et al. [37] demonstrated through a target search task that gaze-based selection significantly reduces target acquisition time compared to using a mouse pointer. In VR selection tasks, Sidenmark et al. [45] found that gaze-based aiming outperforms raycasting selection with a handheld controller, and both methods are superior to target pointing based on head orientation. Although gaze-based interaction shows clear advantages in terms of temporal efficiency, this benefit does not extend to error rates [33]. Scott et al. [43] attributed the higher error rates to intrinsic characteristics of the gaze system: unlike cursor manipulation, *users cannot rely on real-time visual feedback for calibration and must instead predict and make anticipatory adjustments*. During the correction phase, users alternate between fixations and saccades [14], maintaining prolonged fixations (ranging from 100 milliseconds to several seconds) to stabilize the cursor's position while employing saccades to compensate for positioning errors [38].

Research on path-steering tasks in VR employing gaze-based interaction remains scarce. Although the influence of gaze-based interaction has been modeled by an improved Fitts' law for selection tasks [62], it is still unclear whether the insensitivity to distance and the limited utilization of target width observed in gaze-based target selection also manifests in path-steering tasks. Given the differences between target selection and steering tasks, further investigation is needed to understand the efficiency and underlying mechanisms of gaze-based steering in VR.

2.3 Eye Tracking Accuracy and Precision

In 1981, McConkie et al. [29] emphasized the importance of defining and reporting data quality metrics in eye-tracking experiments. Data quality is critical in ensuring the comparability and standardization of experimental results [3, 35]. In line with this, Holmqvist et al. [15] proposed a minimal reporting guideline for eye-tracking research, which mandates that the quality of the fixation signals provided by an eye tracker—such as spatial accuracy, spatial precision, packet loss, and sampling rate—should be reported.

Because current eye tracking systems have achieved substantial control over packet loss and sampling rate—thereby minimizing the impact of external factors [46]—our research focuses primarily on two performance measures: *spatial accuracy* (*Ac*) and *spatial precision* (*Pr*). Spatial accuracy [12] is defined as the mean angular deviation between the recorded fixation points and their corresponding target positions, computed over multiple samples as shown in Eq. (5). In practical terms, a lower average angular deviation indicates that the eye tracker produces measurements that are closer to the true target. Thus, the accuracy is defined as Eq. (5), with g and t_i representing the measured gaze and target vectors, respectively, and N the number of samples.

$$\text{Accuracy} = \frac{1}{N}\sum_{i=1}^{N} \arccos\left(\frac{g_i \cdot t_i}{\|g_i\| \, \|t_i\|}\right) \tag{5}$$

Fernandes et al. [12] artificially introduced additional spatial accuracy offsets to investigate the effects of degraded accuracy on selection tasks. Their findings showed that when spatial accuracy errors were high, users needed more time to complete tasks and made more errors. The new x and y coordinates (in degrees) were derived from the original Tobii gaze point $\left(P_{\text{original}}\right)$ at the same z-distance, yielding a degraded point $\left(P_{\text{degraded}}\right)$ through a rotation θ and offset magnitude Mag.

$$P_{\text{degraded}} = P_{\text{original}} + \text{Vec3}\big(\cos(\theta), \sin(\theta), 0\big) \times \text{Mag} \tag{6}$$

Spatial precision quantifies the stability of individual gaze samples over time for a fixed target [34] and is typically determined by calculating the standard deviation of the angular differences between each individual gaze direction and the mean one. Similarly, the precision is formulated as Eq. (7), where $\bar{g}$ denotes the mean gaze vector computed over all N samples.

$$\text{Precision} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} \|g_i - \bar{g}\|^2} \tag{7}$$

Schuetz et al. [42] introduced Gaussian noise to simulate the random errors caused by insufficient spatial precision in tracking. Their findings demonstrated that spatial accuracy significantly impacts error rates and helped determine the minimum target size on the interface required to ensure smooth user selection. The following equation expresses how Gaussian noise is added to the original gaze position to obtain the degraded position:

$$\mathbf{P}_{\text{degraded}} = \mathbf{P}_{\text{original}} + \begin{pmatrix} \varepsilon_x \\ \varepsilon_y \\ 0 \end{pmatrix}, \quad \varepsilon_x, \varepsilon_y \sim \mathcal{N}(0, \sigma^2) \tag{8}$$

While manufacturers rarely disclose spatial precision and accuracy metrics for their eye-tracking systems, studies have measured these parameters for several popular devices under controlled conditions. These measurements (see Tab. 1) ensure reliable user experiments.

Device	Accuracy(°)	Precision(°)	
		RMS	SD
Apple Vision Pro [18]	2.5	Not stated	Not stated
Quest Pro [49]	2.162	0.772	0.673
HoloLens 2 [5]	2.92	0.071	Not stated
Vive Pro Eye [4]	1.08	0.20	0.36

Table 1: Measured spatial accuracy and precision of commercial eye-tracking devices. Precision is characterized by RMS and SD values.

3 Research Questions

While prior work has extensively explored the modeling of gaze-based input, most studies have focused exclusively on selection tasks, largely overlooking steering interactions and the influence of tracking quality. Thus, to systematically understand how gaze characteristics and path constraints influence *user behaviors with gaze-based steering interaction*, we pose the following research questions:

RQ1: *How do path features, such as width and length, influence user behavior in gaze-based steering tasks?* Previous studies on gaze-based pointing indicate that movement time is not strictly proportional to the amplitude of the movement [13, 41], implying limited sensitivity to path length (A) in selection tasks. Zhang et al. [62] adapted Fitts' Law by applying an exponential transformation to A and reducing the effective target width (W) using an empirically derived constant to better accommodate eye-fixation stability. Inspired by these works, although prior work has examined gaze-based steering performance, it is necessary to revisit this under varied path settings to ensure model accuracy as a foundation for incorporating tracking quality. We therefore aim to investigate how path width and length affect movement time, average steering speed, and error rates in gaze-based tasks (Sec. 4.7).

RQ2: *How do spatial precision and accuracy of eye tracking influence user behavior in gaze-based steering tasks?* As discussed in Sec. 2.3, previous studies have attributed poor task performance in gaze-based interaction primarily to limited eye-tracking quality. To investigate this, researchers have introduced artificial gaze cursor offsets to manipulate spatial accuracy and applied Gaussian noise to degrade spatial precision, providing controlled hardware and environments to simulate suboptimal conditions. While such manipulations significantly affect task completion time and error rates in selection tasks, the distinct nature of steering tasks requires dedicated exploration. This RQ aims to investigate how variations in spatial accuracy and precision influence user behavior in gaze-based steering, focusing on quantitatively characterizing their impact on performance metrics.

RQ3: *How does user behavior in gaze-based cursor control align with the Steering Law?* Unlike mouse-driven cursor movement, eye movements alternate between fixations and saccades [14, 38], which limits the ability to fully utilize visual feedback for real-time control and necessitates predictive adjustments [41]. In addition, inherent ocular behaviors, such as tremors, drifts, and microsaccades, further affect gaze stability and control. This RQ investigates whether the Steering Law, which models the relationship between path characteristics and movement time, remains applicable for gaze-based steering tasks.

RQ4: *Can the effects of spatial accuracy and precision on user behavior in steering tasks be captured or explained through mathematical modeling?* Previous research, see Sec. 2.1, has shown that integrating additional variables—such as latency and frame rate—into traditional interaction models significantly enhances their applicability. Building on insights from ***RQ1*** and ***RQ2*** regarding path characteristics and eye-tracking data quality, and considering the validity of the Steering Law addressed in ***RQ3***, we seek to formalize these relationships into novel predictive models. Such models can explicitly incorporate spatial accuracy and precision as parameters, potentially improving the predictive power and generalizability of models for gaze-based steering interaction.

4 User Study

This study systematically explores how spatial accuracy, spatial precision, and path characteristics (length and width) affect user performance and behavior during gaze-based steering tasks. This research was reviewed and approved by the University Ethics Committee of Xi'an Jiaotong-Liverpool University (ER-LRR-12889407202310231340 13).

4.1 Participants

Twenty-four participants (13 males and 11 females) were recruited from a local university. Their ages ranged from 18 to 26 years ($M = 21.56, SD = 1.59$). We also recorded each participant's visual characteristics and whether they wore corrective lenses (see the supplementary material). Using a 7-point Likert scale, participants rated their familiarity with VR systems ($M = 4.42, SD = 1.28$) and their familiarity with gaze-based interaction ($M = 3.13, SD = 1.90$), with higher scores indicating greater familiarity. The initial spatial precision and accuracy of the users were evaluated following the procedure detailed in Sec. 4.3. The results indicated an initial spatial accuracy of ($M = 0.77°, S = 0.28°$) and a spatial precision of ($M = 0.38°, S = 0.29°$), see Sec. 4.3 with a Meta Quest Pro.

4.2 Apparatus

Eye movements were recorded using the proprietary, integrated eye tracker in the Meta Quest Pro VR headset, which had a resolution of 1800 × 1920 per eye, a horizontal field of view (FoV) of 106°, and a vertical FoV of 95.57°. The eye-tracking hardware, a standard component of the Quest Pro, captures positional gaze data at up to 90 Hz via Meta's Unity public eye-tracking API [31]. To minimize extraneous head movement even further [6] and to guarantee that the cursor is operated only by eye movements, participants' heads were stabilized using a chin rest, similar to those used in vision assessments. The software ran on a PC equipped with an Intel Core i9 processor and an NVIDIA RTX 3080 Ti GPU.

4.3 Initial Spatial Accuracy and Precision Measurement

Before the formal experiment, participants adjusted the interpupillary distance (IPD) on the Quest Pro headset until the display was clearly visible. Subsequently, they completed the 9-point eye-tracking calibration provided by the Quest Pro and self-assessed its calibration quality before proceeding. To quantify eye-tracking quality more precisely, we followed the approach proposed by Wei et al. [49], positioning 13 targets evenly distributed across a $\pm 15°$ field of view at intervals of 5°. Each target had an angular size of approximately 0.7° and was placed at a fixed depth of 0.5 meters from the participants. At the beginning of each trial, participants initiated the task by pressing the trigger button on the handheld controller. Targets appeared red and turned blue upon gaze activation. Participants were instructed to maintain fixation on each target for 5 seconds, during which gaze endpoints were sampled at 90 Hz. Spatial accuracy and precision were then calculated using equations Eq. (5) and Eq. (7), as described in Section 2.3.

4.4 Experimental Task

In our experimental setup, we adopted a task similar to the steering task described in prior VR research [23, 50]. The starting region was marked in green and the ending region in blue [55]. Before each trial, participants pressed a button on the controller to activate the eye-tracking system and control the cursor ball. When the cursor ball overlapped with the starting region for 100 ms, that region turned red, signaling the beginning of the task. The participants were then required to control the cursor ball to move ("push") a target ball with a diameter equal to the path width along the entire path until it reached the blue end area, marking the completion of a trial. If the cursor ball exited the path, we considered the trial to be erroneous.

4.5 Design and Procedure

The user study employed a $2 \times 3 \times 4 \times 4$ within-subjects design with four independent variables, resulting in 96 conditions in total:

- Path Length (A): 30°, and 50°
- Path Width (W): 3° , 4.5° , and 6°
- Artificial Spatial Accuracy (Ac): 0° , 1° , 2° , and 3°
- Artificial Spatial Precision (Pr): 0° , 0.5° , 1° , and 1.5°

Path length was defined as the distance from the starting region to the ending region, while path width referred to the diameter of the target ball (see Fig. 2). The design of path width and path length followed the difficulty index (IDs) from the original Steering Law Eq. (3), ranging from 5.0 to 16.67 bits, covering a spectrum from easy to difficult task conditions.

Since spatial precision and accuracy vary significantly across users and device conditions, we first reviewed existing measurements reported in prior studies (see Tab. 1), where spatial accuracy ranged from 1.08° to 2.92°, and spatial precision (SD) ranged from 0.36° to 0.67°. Inspired by Yamanaka et al.'s approach [56] to analyzing latency effects, we categorized spatial precision and accuracy into baseline values determined by device configurations and additional manipulations intentionally introduced in our study. Specifically, following Fernandes et al. [12], we artificially introduced spatial accuracy offsets (see Eq. (6)), with randomized direction varying across conditions. To simulate the random errors caused by insufficient spatial precision, we adopted the Gaussian noise approach proposed by Schuetz et al. [42] (see Eq. (8)). Based on previous empirical findings, we selected spatial accuracy offsets ranging from 0° to 1.5° and spatial precision values from 0° to 3°, effectively encompassing the majority of real-world scenarios.

The two A, three W, four Ac, and four Pr values were combined into 96 conditions and presented in a randomized order during the experiment. Each condition was repeated four times, resulting in a total of 9216 data points ($2A \times 3W \times 4AC \times 4Pr \times 4$ *repetitions* × 24 *participants*).

Participants first completed a demographic questionnaire, followed by the initial measurement of spatial precision and accuracy (see Section 4.3). After this measurement, participants engaged in a five-minute practice session to familiarize themselves with the task. Formal trials

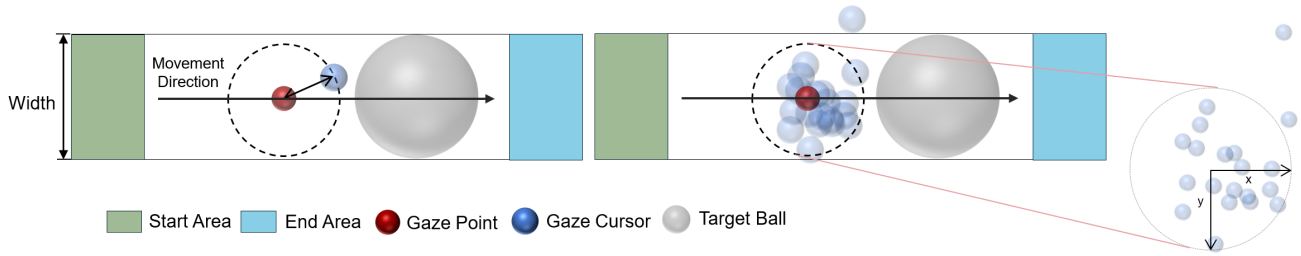

Fig. 2: The path width is defined by the diameter of the target ball. Participants control the gaze cursor along a predefined movement direction to push the target ball from the starting area to the target area. The left illustration depicts a fixed offset representing an intentionally injected spatial accuracy error, while the right illustration shows a spatial precision error simulated using Gaussian noise.

began after this practice period. During the experiment, participants were instructed to remain seated in a fixed, non-rotating chair, minimize head movement, and focus on speed and accuracy. Each session lasted approximately 40 minutes, with short breaks between trials to prevent eye fatigue.

4.6 Evaluation Metrics

Three performance metrics—movement time (MT), average speed (V), and error rate—were collected for each trial.

- **Movement Time (MT):** The total time taken to complete the steering task from the start area to the end region. MT is a standard metric for assessing human behavior patterns and task performance [54].
- **Average Speed (V):** Calculated as the path length divided by movement time, average speed offers a stable performance measure by minimizing the variability seen in point-based speed measurements [59]. Higher average speeds indicate better performance [53].
- **Error Rate:** We recorded a successful completion when the user managed to control the target ball to the endpoint in a single attempt. An error occurred when the cursor either lost contact with the target ball or exited the path before reaching the goal.

4.7 Results

We initially collected 9216 records. After excluding 13 trials (0.14%) with movement times exceeding 20 seconds and 166 trials (1.80%) that deviated more than three standard deviations from the mean speed, we analyzed 9037 valid records. Repeated measures ANOVA was conducted on Movement Time (MT), Average Speed (V), and Error Rate, applying Greenhouse-Geisser correction for sphericity violations and Bonferroni corrections for post-hoc comparisons.

4.7.1 Movement Time

Factors A, W, Ac, and Pr significantly affected movement time (A: $F_{1,23} = 286.067$, $p < 0.001$, $\eta_p^2 = 0.926$; W: $F_{1.088,25.031} = 219.907$, $p < 0.001$, $\eta_p^2 = 0.905$; Ac: $F_{2.391,54.989} = 93.506$, $p < 0.001$, $\eta_p^2 = 0.803$; Pr: $F_{1.332,30.635} = 37.823$, $p < 0.001$, $\eta_p^2 = 0.622$). The interactions were also significant ($A \times W$: $F_{1.416,32.566} = 84.023$, $p < 0.001$, $\eta_p^2 = 0.785$; $W \times Ac$: $F_{6,138} = 6.032$, $p < 0.001$, $\eta_p^2 = 0.208$; $A \times Pr$: $F_{3,69} = 20.471$, $p < 0.001$, $\eta_p^2 = 0.471$; $W \times Pr$: $F_{3.37,77.511} = 4.267$, $p = 0.001$, $\eta_p^2 = 0.156$).

Post-hoc analyses revealed that within factor A, the 30° condition yielded significantly shorter movement times than 50° ($\Delta = 1790$ ms, $p < 0.001$). For factor W, the 3° condition resulted in significantly longer movement times than both 3.5° ($\Delta = 2080$ ms, $p < 0.001$) and 6° ($\Delta = 2817$ ms, $p < 0.001$), while 4.5° was significantly longer than 6° ($\Delta = 737$ ms, $p < 0.001$). Regarding factor Ac, movement time at 0° was significantly shorter than at 2° ($\Delta = 661$ ms, $p < 0.001$) and 3° ($\Delta = 1026$ ms, $p < 0.001$); similarly, 1° was shorter than both 2° ($\Delta = 484$ ms, $p < 0.001$) and 3° ($\Delta = 849$ ms, $p < 0.001$), and 2° was shorter than 3° ($\Delta = 365$ ms, $p < 0.001$). For factor Pr, the 0° condition produced significantly shorter times compared to 1° ($\Delta = 915$ ms, $p < 0.001$) and 1.5° ($\Delta = 1394$ ms, $p < 0.001$); similarly, 0.5° was shorter than 1° ($\Delta = 595$ ms, $p < 0.001$) and 1.5° ($\Delta = 1074$ ms, $p < 0.001$), while 1° was shorter than 1.5° ($\Delta = 479$ ms, $p < 0.001$).

4.7.2 Speed

Factors W, Ac, and Pr also had a significant influence on speed (W: $F_{1.192,27.419} = 178.235$, $p < 0.001$, $\eta_p^2 = 0.886$; Ac: $F_{2.276,52.346} = 65.198$, $p < 0.001$, $\eta_p^2 = 0.739$; Pr: $F_{1.345,30.942} = 48.804$, $p < 0.001$, $\eta_p^2 = 0.680$). Significant interactions were observed as well ($A \times W$: $F_{2,46} = 7.706$, $p = 0.010$, $\eta_p^2 = 0.251$; $W \times Ac$: $F_{6,138} = 89.296$, $p = 0.010$, $\eta_p^2 = 0.200$; $W \times Pr$: $F_{3.702,85.148} = 151.024$, $p = 0.030$, $\eta_p^2 = 0.162$; $Ac \times Pr$: $F_{4.615,106.147} = 151.024$, $p < 0.010$, $\eta_p^2 = 0.192$).

Post-hoc comparisons showed that when W was 3°, speed was significantly lower than at 4.5° ($\Delta = 6.266$ °/s, $p < 0.001$) and 6° ($\Delta = 10.687$ °/s, $p < 0.001$), and the 4.5° condition was lower than 6° ($\Delta = 4.421$ °/s, $p < 0.001$). For factor Ac, speed at 0° was significantly lower than at 2° ($\Delta = 2.991$ °/s, $p < 0.001$) and 3° ($\Delta = 4.126$ °/s, $p < 0.001$); likewise, 1° was lower than 2° ($\Delta = 2.144$ °/s, $p < 0.001$) and 3° ($\Delta = 3.279$ °/s, $p < 0.001$), and 2° was lower than 3° ($\Delta = 1.135$ °/s, $p < 0.001$). Regarding Pr, speed at 0° was significantly lower than at 0.5° ($\Delta = 3.404$ °/s, $p = 0.003$), 1° ($\Delta = 6.733$ °/s, $p < 0.001$), and 1.5° ($\Delta = 8.692$ °/s, $p < 0.001$); similarly, 0.5° was lower than 1° ($\Delta = 3.329$ °/s, $p < 0.001$) and 1.5° ($\Delta = 5.288$ °/s, $p < 0.001$), while 1° was lower than 1.5° ($\Delta = 1.959$ °/s, $p < 0.001$).

4.7.3 Error Rate

Factors W, Ac, and Pr also had a significant influence on speed (W: $F_{1.192,27.419} = 178.235$, $p < 0.001$, $\eta_p^2 = 0.886$; Ac: $F_{2.276,52.346} = 65.198$, $p < 0.001$, $\eta_p^2 = 0.739$; Pr: $F_{1.345,30.942} = 48.804$, $p < 0.001$, $\eta_p^2 = 0.680$). Significant interactions were observed as well ($A \times W$: $F_{2,46} = 7.706$, $p = 0.010$, $\eta_p^2 = 0.251$; $W \times Ac$: $F_{6,138} = 89.296$, $p = 0.010$, $\eta_p^2 = 0.200$; $W \times Pr$: $F_{3.702,85.148} = 151.024$, $p = 0.030$, $\eta_p^2 = 0.162$; $Ac \times Pr$: $F_{4.615,106.147} = 151.024$, $p < 0.010$, $\eta_p^2 = 0.192$).

Post-hoc comparisons showed that when W was 3°, speed was significantly lower than at 4.5° ($\Delta = 6.266$ °/s, $p < 0.001$) and 6° ($\Delta = 10.687$ °/s, $p < 0.001$), and the 4.5° condition was lower than 6° ($\Delta = 4.421$ °/s, $p < 0.001$). For factor Ac, speed at 0° was significantly lower than at 2° ($\Delta = 2.991$ °/s, $p < 0.001$) and 3° ($\Delta = 4.126$ °/s, $p < 0.001$); likewise, 1° was lower than 2° ($\Delta = 2.144$ °/s, $p < 0.001$) and 3° ($\Delta = 3.279$ °/s, $p < 0.001$), and 2° was lower than 3° ($\Delta = 1.135$ °/s, $p < 0.001$). Regarding Pr, speed at 0° was significantly lower than at 0.5° ($\Delta = 3.404$ °/s, $p = 0.003$), 1° ($\Delta = 6.733$ °/s, $p < 0.001$), and 1.5° ($\Delta = 8.692$ °/s, $p < 0.001$); similarly, 0.5° was lower than 1° ($\Delta = 3.329$ °/s, $p < 0.001$) and 1.5° ($\Delta = 5.288$ °/s, $p < 0.001$), while 1° was lower than 1.5° ($\Delta = 1.959$ °/s, $p < 0.001$).

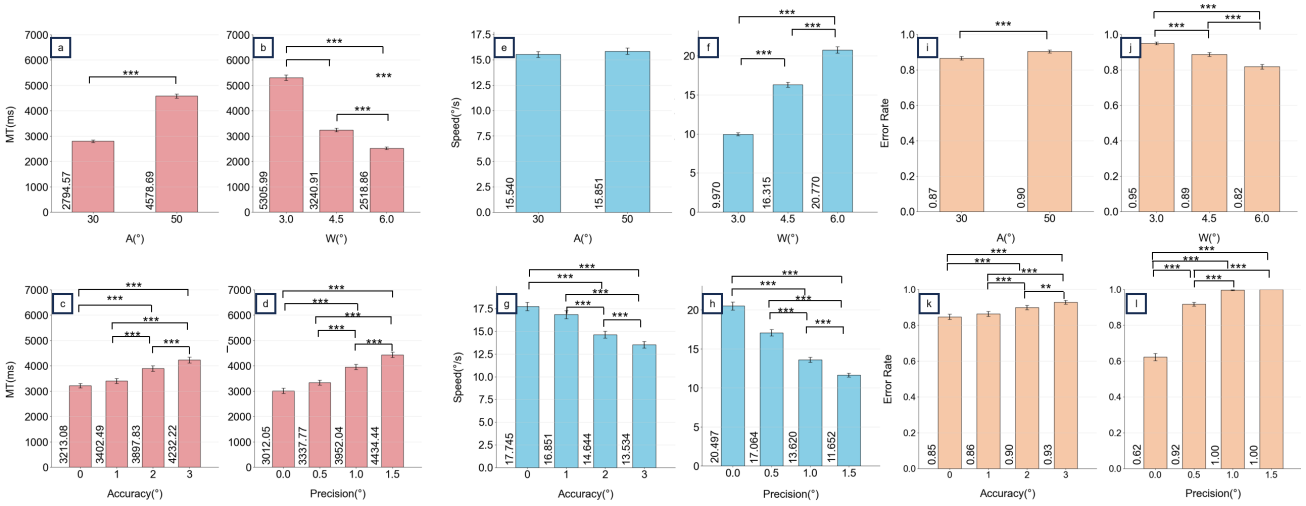

Fig. 3: The effects of path length (A), the path width (W), spatial accuracy (Ac), and spatial precision (Pr) on (a–d) movement time, (e–h) speed, and (i–l) error rates. Error bars represent 95% confidence intervals. Significance levels are indicated as follows: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

4.8 Discussion

4.8.1 Effects of Path Features in Gaze-Based Steering Tasks (*RQ1*)

Our results demonstrated that both path amplitude (A) and width (W) significantly influenced movement time and error rate ($p < 0.001$) for gaze-based steering, aligned with prior findings from traditional steering studies [55]. Regarding cursor speed, path width had a significant impact, whereas amplitude did not. Interestingly, despite the wide range of difficulty indices selected for this study, no speed saturation effect was observed, potentially due to unique constraints inherent to gaze-based interaction [57]. Notably, the error rates recorded in our gaze-based steering tasks were substantially higher than those reported in studies employing hand, controller, or mouse inputs [50, 52]. We attribute this elevated error rate primarily to two factors. First, inherent characteristics of gaze interaction—such as the inability of the eyes to fully exploit the target width and uncontrolled micro-saccadic movements—typically lead to increased error rates [61]. Second, limitations in spatial accuracy and precision inherent to eye-tracking hardware introduce cursor instability, further compromising users' ability to effectively control the cursor, thus exacerbating error rates.

4.8.2 Effects of Spatial Accuracy and Precision in Gaze-Based Steering Tasks (*RQ2*)

Our analysis revealed that both *spatial accuracy* (Ac) and *spatial precision* (Pr) significantly impacted movement time (see Fig. 3). Post-hoc pairwise comparisons showed that at lower levels of spatial accuracy error ($Ac = 0°, 1°$) and spatial precision error ($Pr = 0°, 0.5°$), no statistically significant differences in user performance were observed. However, increasing spatial accuracy and precision errors substantially impacted movement time, cursor speed, and error rates. Notably, error rates approached 100%, indicating saturation at higher levels of spatial precision errors ($Pr = 1.0°, 1.5°$).

Participants generally adapted to spatial accuracy errors by compensating for systematic biases during the initial phases of each task, albeit at the cost of increased initial reaction and adjustment time. In contrast, spatial precision errors presented a more critical challenge due to unpredictable cursor fluctuations, significantly reducing users' ability to anticipate cursor positions from one moment to the next. This unpredictability caused confusion and significantly elevated error rates. Unlike the systematic adaptation observed for spatial accuracy, participants could not effectively adjust to spatial precision errors, resorting instead to maintaining the cursor near the path center as a mitigation strategy.

5 Modeling

In this section, we aim to (1) assess the applicability of the Steering Law in gaze-based interaction, (2) use regression to analyze the impact of spatial accuracy and spatial precision on time, (3) develop new candidate models based on generalized eta-squared and eye movement characteristics, and (4) compare the predictive performance of the candidate models with the original Steering Law.

5.1 Verifying the Applicability of the Steering Law in Gaze-Based Interaction

First, we independently analyze spatial precision and spatial accuracy to isolate their effects and examine the original Steering Law formulation in the context of gaze-based control for the Steering Task. This allows us to assess the predictive accuracy of movement time (MT) under this control modality across different path lengths (A) and path widths (W).

A total of 16 combinations ($4Ac \times 4Pr$) were considered. Within each $Ac \times Pr$ combination, there were six unique IDs ($2A \times 3W$). We applied Ordinary Least Squares (OLS) regression separately to each of the 16 combinations. Fig. 4 shows the fitting results obtained using the original Steering Law applied to all variables ($N = 96$; $2A \times 3W \times 4Ac \times 4Pr$), resulting in an adjusted $R^2 = 0.798$. Subsequently, we fitted the model separately for each spatial precision and accuracy combination, resulting in 6 data points ($2A \times 3W$) per combination. The original Steering Law Eq. (3) demonstrated excellent fits, as indicated by the coefficient of determination ($adj.R^2$) for each $Ac \times Pr$ combination (M = 0.976, SD = 0.021).

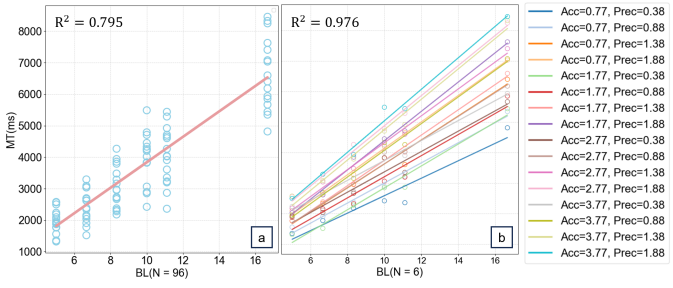

Fig. 4: (a) Model fitting of movement time (MT) using the original Steering Law across all conditions ($n = 90$). (b) Separate model fittings for each combination of spatial precision and spatial accuracy.

5.2 Formulating the Effect of Spatial Accuracy and Precision

In this section, we employ regression analysis to mathematically quantify the complex relationship between spatial precision (Ac) and spatial accuracy (Pr) with movement time (MT). In Sec. 4.7.1, we report the effects of spatial precision and accuracy on movement time, as well as their interaction effects with path width and path length. Our findings indicate that when spatial precision and accuracy are low, their impact is not significant. However, at higher levels of spatial precision and accuracy, their influence becomes more pronounced. Based on this observed trend, we explored four commonly used regression models: *linear, power, logarithmic, and exponential functions*. To evaluate the fitting performance of these regression functions and prevent overfitting, we computed the coefficient of determination (R^2), the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC). As shown in Sec. 5.2, both partial precision and accuracy exhibited the best fit with the linear and exponential functions. Furthermore, since the ΔAIC and ΔBIC values for these two functions are less than 2, their performance differences are not considered statistically significant [8, 40]. To minimize the complexity introduced by additional variables in subsequent modeling, we selected the linear function to quantify the effect of spatial precision and accuracy on movement time in future analyses.

Factor	Model	Adjusted R^2	AIC	BIC
Ac	$MT = a + bAc$	**0.975**	**37.214**	**35.986**
	$MT = a + b \times Ac^c$	0.902	42.686	41.45
	$MT = a + b\ln(Ac)$	0.877	43.580	42.353
	$MT = a + b \times e^{Ac}$	**0.980**	**36.189**	**34.962**
Pr	$MT = a + bPr$	**0.987**	**36.911**	**35.684**
	$MT = a + b \times Pr^c$	0.931	43.766	42.539
	$MT = a + b\ln(Pr)$	0.899	45.264	44.036
	$MT = a + b \times e^{Pr}$	**0.992**	**35.042**	**33.815**

Table 2: Comparison of four regression models (linear, power, logarithmic, and exponential) for quantifying the relationship between spatial accuracy (Ac), spatial precision (Pr), and movement time (MT). Best-performing results are highlighted.

5.3 Model Derivation

Given the absence of prior work specifically addressing gaze-based steering tasks, we first adopted Zhang et al.'s extension of Fitts' Law (Eq. (4)) [62], which integrates gaze-specific characteristics for selection tasks, as our **Candidate Model 1 (CM1)**. To determine whether spatial accuracy and precision should be incorporated into predictive steering models, we computed generalized eta-squared (η_G^2) values to quantify their contributions to movement time variance. The results showed substantial effect sizes for amplitude (A, $\eta_G^2 = 0.372$), width (W, $\eta_G^2 = 0.512$), accuracy (Ac, $\eta_G^2 = 0.108$), precision (Pr, $\eta_G^2 = 0.176$), and the interaction between amplitude and width ($A \times W$, $\eta_G^2 = 0.076$). Following Cohen's criteria for effect sizes [9] (small: 0.01, medium: 0.06, large: 0.14), only factors and interactions with medium or larger effects were considered further. Our findings clearly indicate that both spatial accuracy and precision significantly influence steering performance without notable interaction effects. Thus, incorporating these two gaze-tracking quality factors into the predictive models is warranted.

Based on these quantified relationships (see Section Sec. 5.2), we extended the original Steering Law (Eq. (3)) by linearly integrating spatial accuracy to create **Candidate Model 2 (CM2)**. Similarly, spatial precision was incorporated into the original model separately, resulting in **Candidate Model 3 (CM3)**. Finally, both spatial accuracy and precision were simultaneously integrated into the original model to develop **Candidate Model 4 (CM4)**. The detailed derivation and structure of these candidate models are summarized in Sec. 5.4.2.

5.4 Model Evaluation

5.4.1 Model Fitting Methods and Metrics

We used non-linear least squares optimization to accurately estimate parameters for each candidate model. Multiple initial parameter values were tested during fitting to ensure convergence toward the global optimum. The final model was selected based on achieving the highest adjusted R^2 and the lowest AIC and BIC values, as shown in Sec. 5.2. Given the comparable performance of the linear and exponential models ($\Delta AIC, \Delta BIC < 2$), we opted for the linear function due to its interpretability and simplicity in subsequent modeling.

5.4.2 Evaluation Results

Sec. 5.4.2 summarizes the nonlinear regression analysis results, comparing the baseline Steering Law model against four candidate models. Among these, CM4, integrating both spatial accuracy and precision, yielded the best predictive performance, clearly indicated by substantial improvements in information criteria (ΔAIC > 10, ΔBIC > 10). CM3, incorporating only spatial accuracy, ranked second, closely followed by CM2, which considered only spatial precision. Both CM3 and CM2 significantly outperformed the baseline in terms of adjusted R^2, AIC, and BIC. Conversely, CM1 demonstrated negligible improvement relative to the baseline, emphasizing the necessity of incorporating both spatial accuracy and precision to substantially enhance predictive accuracy.

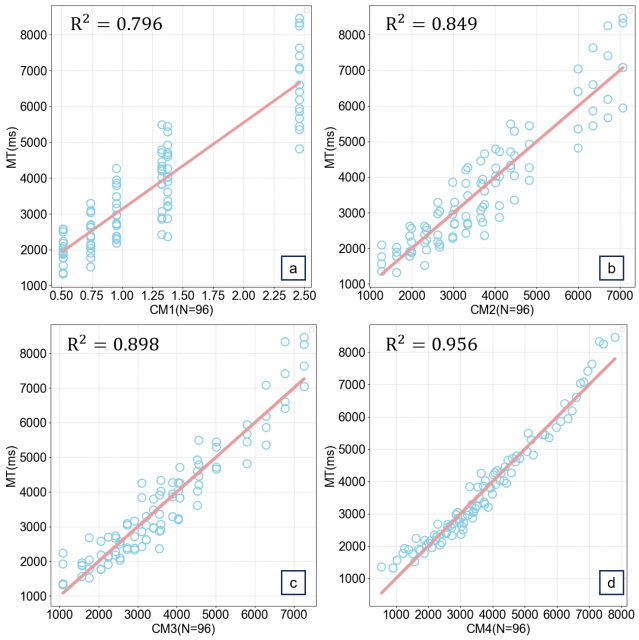

Fig. 5: Comparison of model fitting performance for four candidate models (CM1–CM4), across all conditions (N = 96)

5.4.3 Cross-Validation

We conducted cross-validation tests based on condition grouping to verify the generalizability of the four models. Model coefficients were obtained from 72 randomly selected experimental conditions (levels), and the model fit was tested on the remaining 24 conditions over 100 iterations. Specifically, the baseline model achieved an average R^2 of 0.7592 ($\pm$0.0917) and an MSE of 603893 ($\pm$163822). Among the four candidate models, Model 1 obtained an average R^2 of 0.7621 ($\pm$0.0765) and an MSE of 620684 ($\pm$148607); Model 2 achieved an average R^2 of 0.8156 ($\pm$0.0589) and an MSE of 448891 ($\pm$106534); Model 3 recorded an average R^2 of 0.8776 ($\pm$0.0412) and an MSE of 308981 ($\pm$80890). Notably, Model 4 demonstrated the highest

Model	R^2	AIC	BIC
BL: Eq. (3)	0.795	1550	1555
CM1: Eq. (4)	0.796	1278	1288
CM2: $MT = a + b\frac{A}{W} + c \cdot Ac$	0.849	1248	1256
CM3: $MT = a + b\frac{A}{W} + c \cdot Pr$	0.898	1211	1218
CM4: $MT = a + b\frac{A}{W} + c \cdot Ac + d \cdot Pr$	**0.956**	**1132**	**1142**

Table 3: Fitting results, where 'BL' denotes the baseline and 'CM' represents the candidate models. The best performance results are highlighted and in bold. Each candidate model incrementally adds predictors to the previous one, i.e., they are nested, and improvements are clearly reflected in the decreasing AIC and BIC values.

predictive power, with an average R^2 of 0.9496 ($\pm$0.0136) and an MSE of 138727 ($\pm$35424). Overall, our findings indicate that all models achieved accurate predictions in the cross-validation analysis, which were also similar to the original estimates. This suggests that the models are capable of predicting unseen experimental conditions with high accuracy.

5.5 Discussion

5.5.1 The Applicability of Steering Law in Gaze-Based Cursor Control Task (*RQ3*)

Our findings revealed that fitting the original Steering Law (Eq. (3)) separately for each combination of spatial accuracy and precision yielded a high adjusted $R^2 = 0.956$. This suggests that, despite the increased complexity of transitioning from mouse-based to gaze-based steering tasks, the Steering Law retains robust explanatory power regarding user performance and behavior.

5.5.2 Modeling User's Behavior in Gaze-Based Steering Tasks (*RQ4*)

We first modeled the relationship between spatial accuracy, spatial precision, and movement time by fitting common regression functions (see Sec. 5.2). Based on model selection criteria (adjusted R^2, AIC, and BIC), a linear relationship provided the best representation of the effects of these two variables on movement time. Subsequently, generalized eta-squared values were analyzed to determine which variables should be incorporated into the original Steering Law. Spatial accuracy ($\eta^2_G = 0.108$) and spatial precision ($\eta^2_G = 0.176$) emerged as the significant factors. Although both parameters reflect the quality of eye-tracking devices, their interaction effects were negligible ($\eta^2_G < 0.01$), justifying their independent incorporation into the predictive models.

Consequently, we constructed three extended models: Candidate Model 2 (incorporating spatial accuracy), Candidate Model 3 (incorporating spatial precision), and Candidate Model 4 (simultaneously incorporating both variables), as detailed in Sec. 5.4.2. We then evaluated these models against the baseline (original Steering Law) and Candidate Model 1 proposed by Zhang et al. Eq. (4). Results demonstrated that our extended models substantially improved predictive performance, as evidenced by adjusted (R^2), AIC, and BIC scores (see Sec. 5.4.2).

Although Candidate Model 1, which accounts for gaze-specific characteristics, showed minimal improvement in adjusted (R^2) compared to the baseline (an increase of only 0.1%), it significantly reduced AIC (by 272) and BIC (by 267), suggesting enhanced parsimony. Candidate Model 4, integrating both spatial accuracy and precision, demonstrated the strongest predictive performance, achieving a notable increase in adjusted (R^2) of 16.1% compared to the baseline, accompanied by substantial reductions in AIC (418 points) and BIC (413 points). This superior performance can be attributed to its comprehensive coverage of explanatory variables. Furthermore, comparisons between Candidate Models 2 and 3 aligned well with generalized eta-squared results, highlighting that precision (Pr) had a stronger impact on predictive accuracy than accuracy (Ac)), with Candidate Model 3 outperforming Candidate Model 2 (adjusted (R^2) improvement of 4.9%, AIC reduction of 37, BIC reduction of 38).

In summary, the four candidate models proposed here deepen our understanding of user behavior in gaze-based steering tasks. These models effectively capture the influence of eye-tracking quality parameters while maintaining interpretability and simplicity, thus minimizing the risk of overfitting.

6 Model Verification

To further validate the robustness of our proposed models, we conducted additional experiments using a different head-mounted display. Unlike Study 1, which employed the Meta Quest Pro with a proprietary eye-tracking system, this experiment utilized the Pico headset equipped with Tobii eye-tracking and tested multiple sampling rates. Steering tasks were performed under various combinations of path characteristics, spatial accuracy, and spatial precision. Subsequently, we applied the candidate models derived from Study 1 to predict movement time and evaluate their stability across different levels of eye-tracking quality.

6.1 Participants, Apparatus, and Task

Fourteen participants (9 males and 5 females) were recruited from a local university, where five had previously participated in the first experiment. Their ages ranged from 18 to 26 years ($M = 20.3, SD = 1.25$). Eye-related information for each participant, including visual characteristics and whether corrective lenses were worn, is provided in the supplementary material. Participants rated their familiarity with VR systems on a 7-point Likert scale ($M = 4.07, SD = 1.75$) and their familiarity with gaze-based interaction on the same scale ($M = 2.28, SD = 1.32$), with higher values indicating greater familiarity. Initial eye-tracking quality was evaluated using the same procedure as described in Sec. 4.3. The results showed an average spatial accuracy of ($M = 0.59°, SD = 0.11°$) and an average spatial precision of ($M = 0.34°, SD = 0.14°$).

The experiment was conducted using a Pico 4 Pro HMD featuring a resolution of 2160 × 2160 pixels per eye and a maximum refresh rate of 90 Hz. The horizontal field of view (FOV) is 104°, and the vertical FOV is 103°. Eye tracking was performed using the device's built-in cameras and accessed through Pico's official API. The experimental application was developed and run on a desktop PC with specifications identical to those used in the first study.

Following the same procedure as in Experiment 1 (see Sec. 4), the task flow was controlled through defined starting, task, and ending regions. Path parameters were varied in terms of width and length, while eye-tracking quality was manipulated through combinations of spatial accuracy, spatial precision, and sampling rate. Participants were instructed to balance both speed and accuracy during task execution.

6.2 Design and Procedure

To assess the stability of the proposed model, we investigated two sampling rates for eye-tracking devices (30 Hz and 72 Hz) as the main independent variable for this study. A within-subjects design was employed, with two eye-tracking sampling rates S (30 Hz, 72 Hz). To vary the task and to keep the data comparable with the first study, the experimental factors also included two path lengths A (2°, 3°), two path widths W (20°, 30°), three levels of spatial precision Ac (0.5°, 1.5°, 2.5°), and three levels of spatial accuracy Pr (0.25°, 0.5°, 0.75°). The experimental procedure and factor ordering were consistent with those described in the first user study (see Sec. 4). Each participant took approximately 30 minutes to complete the experiment. In total, we collected ($2A \times 2W \times 3Ac \times 3Pr \times 2S \times 4$ repetitions $\times$ 14 participants) = 4032 trials.

6.3 Predicting Average Movement Time

Data were preprocessed using the same procedures as described in the first study. Initially, we collected a total of 4,032 records. After excluding 26 trials (0.64%) with movement times exceeding 20 seconds and 94 trials (2.33%) that deviated more than three standard deviations from the mean speed, we retained 3,912 valid trials for analysis. To assess the robustness and generalizability of our models across varying

device conditions and eye-tracking sampling rates, separate analyses were conducted at sampling rates of 30 Hz and 72 Hz. At a sampling rate of 30 Hz, the **baseline** (BL) model yielded an adjusted R^2 of 0.66, exhibiting the lowest predictive performance across all metrics. In contrast, **Candidate Model 4** (CM4), which incorporates all previously identified variables from Experiment 1, achieved a substantially higher adjusted R^2 of 0.904. At 72 Hz, the BL model again showed limited predictive capability with an adjusted R^2 of 0.68, while CM4 exhibited even greater predictive accuracy ($R^2 = 0.925$). These findings demonstrate the strong generalizability and robustness of CM4 across different sampling rates and hardware conditions.

7 Applications

In this section, we present three example application scenarios. As illustrated in Fig. 6a, a user interacts with a cascading menu using a gaze cursor in VR. Under conditions with lower device eye tracking quality, Candidate Model 4 can be used to adjust the menu's width and height to optimize the user experience. Another example is shown in Fig. 6b, where the user avoids interruptive pop-up notifications through gaze steering. For instance, when a message or alert appears, the user can steer their gaze along a narrow path to slide the message or alert up and out of view. Here, our model can help adjust the width of the control bars based on variations in spatial accuracy and precision across different user groups or eye-tracking systems. Finally, as gaze interaction is the primary input for the Apple Vision Pro, Fig. 6c shows an example scenario in which the user answers a phone call via gaze-steering. By steering the gaze to the left, the user declines the call; by steering to the right, the user accepts it.

As one of the most intuitive forms of interaction in VR, gaze-based interaction offers unique advantages, and more application scenarios leveraging gaze-based steering can be added in the future. We believe that our exploration of spatial accuracy and precision can assist designers in adjusting path parameters based on the performance characteristics of different eye-tracking systems. For instance, designers may personalize adaptive gaze-steering tasks for individual users based on spatial accuracy and precision metrics obtained during eye-tracker calibration if a user's device exhibits spatial accuracy errors below 1° or spatial precision instability remains within 0.5°, as we consider jitter within these thresholds to be acceptable (see Sec. 4.7). When eye-tracking quality is suboptimal, our proposed formula (see Sec. 5.4) can help balance task difficulty across users in gaze steering scenarios like the one presented earlier, ensuring consistent user performance and experience. This approach ultimately mitigates negative user experiences caused by eye-tracking limitations across HMDs.

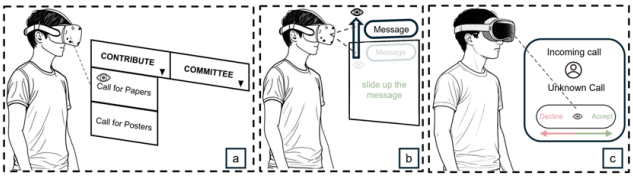

Fig. 6: Application scenarios for gaze-based steering interaction, where (a) illustrates a gaze-controlled cascading menu utilizing steering behavior for item selection, (b) a notification dismissal mechanism, where users steer their gaze along a narrow path to remove pop-ups without manual interaction, (c) presents a scenario on Apple Vision Pro where a phone call is answered using gaze steering.

8 Limitations and Future Work

Our work identified two main limitations, which point to directions for future work. First, while our model demonstrated favorable outcomes in Study 1 through modeling with linear paths and cross-validation, it explored only a subset of all potential paths. Our model's robustness was further substantiated in Study 2 under varying sampling rates and across different devices. However, given the complexity of real-world scenarios, future work needs to explore the impact of additional path characteristics to investigate the interaction effects between gaze-based interaction and a broader range of path parameters. Second, although our findings indicate that limited, small spatial precision ($Ac < 1°$) and accuracy ($Pr < 0.5°$) errors do not significantly impact movement time, understanding the relationship between these interaction characteristics and their respective thresholds is also crucial. Thus, we plan to systematically analyze thresholds across various interaction scenarios in the future and will explore potential dependencies between these thresholds and interaction characteristics. Given the susceptibility of gaze-based input to fatigue, exploring how different levels of task difficulty impact fatigue during gaze-based steering tasks will be important for designing long-duration applications.

9 Conclusion

In this paper, we presented our exploration of gaze-based steering interaction for virtual reality (VR) systems, focusing on the effect of the spatial precision and accuracy of eye-trackers on steering tasks. Two user studies were conducted to explore and model gaze-based steering in VR. The first study collected user behavior data for various path characteristics and eye-tracking conditions. The results enabled us to formulate four novel models, where the best one achieved an adjusted R^2 of 0.956 and delivered a 16% improvement in movement time prediction and 26.9% and 26.5% enhancements in AIC and BIC, respectively. Data from the second study with varied technical settings, such as a different sampling rate and a different VR HMD, helped confirm the robustness and predictability of our models. Overall, our work improves our understanding of the effects of eye-tracker precision and accuracy on gaze-based steering interaction in VR systems.

Acknowledgment

The authors thank the participants for their time and the reviewers for their comments, which helped improve our work.

References

[1] J. Accot and S. Zhai. Beyond fitts' law: models for trajectory-based hci tasks. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI '97, 8 pages, p. 295–302. Association for Computing Machinery, New York, NY, USA, 1997. doi: 10.1145/258549.258760 1, 2

[2] D. Ahlström. Modeling and improving selection in cascading pull-down menus using fitts' law, the steering law and force fields. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 61–70, 2005. 2

[3] D. Akkil, P. Isokoski, J. Kangas, J. Rantala, and R. Raisamo. Traqume: a tool for measuring the gaze tracking quality. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '14, 4 pages, p. 327–330. Association for Computing Machinery, New York, NY, USA, 2014. doi: 10.1145/2578153.2578192 3

[4] Anonymous. Eye tracking in virtual reality: Vive pro eye spatial accuracy, precision, and calibration reliability. *Journal of Eye Movement Research*, 15(3), 2022. 3

[5] S. Aziz and O. Komogortsev. An assessment of the eye tracking signal quality captured in the hololens 2. In *2022 Symposium on eye tracking research and applications*, pp. 1–6, 2022. 3

[6] S. Aziz, D. J. Lohr, L. Friedman, and O. Komogortsev. Evaluation of eye tracking signal quality for virtual reality applications: A case study in the meta quest pro. In *Proceedings of the 2024 Symposium on Eye Tracking Research and Applications*, ETRA '24, article no. 7, 8 pages. Association for Computing Machinery, New York, NY, USA, 2024. doi: 10.1145/3649902.3653347 4

[7] J. Blattgerste, P. Renner, and T. Pfeiffer. 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, article no. 1, 9 pages. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3206343.3206349 2

[8] K. P. Burnham and D. R. Anderson. Multimodel inference: understanding aic and bic in model selection. *Sociological methods & research*, 33(2):261–304, 2004. 7

[9] J. Cohen. *Statistical Power Analysis for the Behavioral Sciences*. Routledge, 2nd ed., 1988. doi: 10.4324/9780203771587 7

[10] A. T. Duchowski. Gaze-based interaction: A 30 year retrospective. *Computers & Graphics*, 73:59–69, 2018. doi: 10.1016/j.cag.2018.04.002 2

[11] A. M. Feit, S. Williams, A. Toledo, A. Paradiso, H. Kulkarni, S. Kane, and M. R. Morris. Toward everyday gaze input: Accuracy and precision of eye tracking and implications for design. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, 13 pages, p. 1118–1130. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3025453.3025599 2

[12] A. S. Fernandes, T. S. Murdison, I. Schuetz, O. Komogortsev, and M. J. Proulx. The effect of degraded eye tracking accuracy on interactions in vr. In *Proceedings of the 2024 Symposium on Eye Tracking Research and Applications*, ETRA '24, article no. 63, 7 pages. Association for Computing Machinery, New York, NY, USA, 2024. doi: 10.1145/3649902.3656369 2, 3, 4

[13] K. Z. Gajos, J. O. Wobbrock, and D. S. Weld. Automatically generating user interfaces adapted to users' motor and vision capabilities. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology*, UIST '07, 10 pages, p. 231–240. Association for Computing Machinery, New York, NY, USA, 2007. doi: 10.1145/1294211.1294253 3

[14] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, J. Halszka, and J. van de Weijer. *Eye Tracking : A Comprehensive Guide to Methods and Measures*. Oxford University Press, United Kingdom, 2011. 3, 4

[15] K. Holmqvist, S. L. Örbom, I. T. Hooge, D. C. Niehorster, Alexander, et al. Retracted article: Eye tracking: empirical foundations for a minimal reporting guideline. *Behavior research methods*, 55(1):364–416, 2023. 3

[16] X. Hu, X. Yan, Y. Wei, W. Xu, Y. Li, Y. Liu, and H.-N. Liang. Exploring the effects of spatial constraints and curvature for 3d piloting in virtual environments. In *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 505–514, 2024. doi: 10.1109/ISMAR62088.2024.00065 1

[17] X. Hu, Y. Zhang, Y. Wei, Y. Li, W. Stuerzlinger, and H.-N. Liang. Exploring and modeling gaze-based steering behavior in virtual reality. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '25, article no. 260, 8 pages. Association for Computing Machinery, New York, NY, USA, 2025. doi: 10.1145/3706599.3720273 3

[18] Z. Huang, G. Zhu, X. Duan, R. Wang, Y. Li, S. Zhang, and Z. Wang. Measuring eye-tracking accuracy and its impact on usability in apple vision pro. *arXiv preprint arXiv:2406.00255*, 2024. 3

[19] T. Kim, A. Ham, S. Ahn, and G. Lee. Lattice menu: A low-error gaze-based marking menu utilizing target-assisted gaze gestures on a lattice of visual anchors. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, article no. 277, 12 pages. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3491102.3501977 1, 3

[20] B. Lange, P. Requejo, S. Flynn, A. Rizzo, F. Valero-Cuevas, L. Baker, and C. Winstein. The potential of virtual reality and gaming to assist successful aging with disability. *Physical Medicine and Rehabilitation Clinics of North America*, 21(2):339–356, 2010. doi: 10.1016/j.pmr.2009.12.007 1

[21] J. Lee, H. Kim, Y. Yang, and G. J. Kim. Rpg: Rotation technique in vr locomotion using peripheral gaze. *Proc. ACM Hum.-Comput. Interact.*, 8(ETRA), article no. 235, 19 pages, May 2024. doi: 10.1145/3655609 3

[22] L. Liu, J.-B. Martens, and R. v. Liere. Revisiting path steering for 3D manipulation tasks. *International Journal of Human-Computer Studies*, 69(3):170–181, 2011. doi: 10.1016/j.ijhcs.2010.11.006 1

[23] L. Liu, J.-B. Martens, and R. van Liere. Revisiting path steering for 3d manipulation tasks. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 39–46, 2010. doi: 10.1109/3DUI.2010.5444724 4

[24] X. Lu, D. Yu, H.-N. Liang, and J. Goncalves. itext: Hands-free text entry on an imaginary keyboard for augmented reality systems. In *The 34th Annual ACM Symposium on User Interface Software and Technology*, UIST '21, 11 pages, p. 815–825. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3472749.3474788 2, 3

[25] X. Lu, D. Yu, H.-N. Liang, W. Xu, Y. Chen, X. Li, and K. Hasan. Exploration of hands-free text entry techniques for virtual reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 344–349, 2020. doi: 10.1109/ISMAR50242.2020.00061 2

[26] J. J. MacInnes, S. Iqbal, J. Pearson, and E. N. Johnson. Wearable eye-tracking for research: Automated dynamic gaze mapping and accuracy/precision comparisons across devices. *BioRxiv*, 2018. 2

[27] I. S. MacKenzie. Fitts' law as a research and design tool in human-computer interaction. *Hum.-Comput. Interact.*, 7(1):91–139, 49 pages, Mar. 1992. doi: 10.1207/s15327051hci0701_3 2

[28] P. Majaranta and A. Bulling. *Eye Tracking and Eye-Based Human–Computer Interaction*, pp. 39–65. Springer London, London, 2014. doi: 10.1007/978-1-4471-6392-3_3 2

[29] G. W. McConkie. Evaluating and reporting data quality in eye movement research. *Behavior Research Methods & Instrumentation*, 13(2):97–106, 1981. doi: 10.3758/BF03207916 3

[30] X. Meng, W. Xu, and H.-N. Liang. An exploration of hands-free text selection for virtual reality head-mounted displays. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 74–81, 2022. doi: 10.1109/ISMAR55827.2022.00021 3

[31] Meta: Eye tracking for movement sdk for unity. https://developers.meta.com/horizon/documentation/unity/move-eye-tracking/, 2024. 4

[32] D. Miniotas. Application of fitts' law to eye gaze interaction. In *CHI '00 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '00, 2 pages, p. 339–340. Association for Computing Machinery, New York, NY, USA, 2000. doi: 10.1145/633292.633496 2

[33] A. K. Mutasim, A. U. Batmaz, and W. Stuerzlinger. 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, article no. 15, 7 pages. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3448018.3457998 3

[34] D. C. Niehorster, R. Zemblys, T. Beelders, and K. Holmqvist. Characterizing gaze position signals and synthesizing noise during fixations in eye-tracking data. *Behavior Research Methods*, 52(6):2515–2534, 2020. Published: 2020/12/01. doi: 10.3758/s13428-020-01400-9 3

[35] M. Nyström, R. Andersson, K. Holmqvist, et al. The influence of calibration method and eye physiology on eyetracking data quality. *Behavior Research Methods*, 45:272–288, 2013. doi: 10.3758/s13428-012-0247-4 3

[36] M. Nyström, R. Andersson, K. Holmqvist, and J. van de Weijer. The influence of calibration method and eye physiology on eyetracking data quality. *Behavior Research Methods*, 45(1):272–288, 2013. doi: 10.3758/s13428-012-0247-4 2

[37] O. Oyckoya and F. Stentiford. A performance comparison of eye tracking and mouse interfaces in a target image identification task. In *The 2nd European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology, 2005. EWIMT 2005. (Ref. No. 2005/11099)*, pp. 139–144, 2005. doi: 10.1049/ic.2005.0723 3

[38] C. Prablanc, D. Massé, and J. Echallier. Error-correcting mechanisms in large saccades. *Vision Research*, 18(5):557–560, 1978. doi: 10.1016/0042-6989(78)90202-X 3, 4

[39] M. Proulx. Gaze inputs for targeting: The eyes have it, not with a cursor. *International Journal of Human-Computer Interaction*, 41(1):1–12, 2025. doi: 10.1080/10447318.2025.2453966 2

[40] A. E. Raftery. Bayesian model selection in social research. *Sociological methodology*, pp. 111–163, 1995. 7

[41] I. Schuetz, T. S. Murdison, K. J. MacKenzie, and M. Zannoli. An explanation of fitts' law-like performance in gaze-based selection tasks using a psychophysics approach. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, 13 pages, p. 1–13. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290605.3300765 3, 4

[42] I. Schuetz, T. S. Murdison, and M. Zannoli. A psychophysics-inspired model of gaze selection performance. In *ACM Symposium on Eye Tracking Research and Applications*, ETRA '20 Short Papers, article no. 25, 5 pages. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3379156.3391336 3, 4

[43] S. H. Scott. Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews Neuroscience*, 5(7):532–545, 2004. doi: 10.1038/nrn1427 3

[44] L. E. Sibert and R. J. K. Jacob. Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '00, 8 pages, p. 281–288. Association for Computing Machinery, New York, NY, USA, 2000. doi: 10.1145/332040.332445 2

[45] L. Sidenmark, F. Prummer, J. Newn, and H. Gellersen. Comparing gaze, head and controller selection of dynamically revealed targets in head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 29(11):4740–4750, 2023. doi: 10.1109/TVCG.2023.3320235 3

[46] Tobii. Tobii pro spectrum. https://www.tobii.com/products/eye-trackers/screen-based/tobii-pro-spectrum, 2025. 3

[47] R. Vertegaal. A fitts law comparison of eye tracking and manual input in

the selection of visual targets. In *Proceedings of the 10th International Conference on Multimodal Interfaces*, ICMI '08, 8 pages, p. 241–248. Association for Computing Machinery, New York, NY, USA, 2008. doi: 10.1145/1452392.1452443 2

[48] O. Špakov, H. Istance, K.-J. Räihä, T. Viitanen, and H. Siirtola. Eye gaze and head gaze in collaborative games. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*, ETRA '19, article no. 85, 9 pages. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3317959.3321489 3

[49] S. Wei, D. Bloemers, and A. Rovira. A preliminary study of the eye tracker in the meta quest pro. In *Proceedings of the 2023 ACM International Conference on Interactive Media Experiences*, IMX '23, 6 pages, p. 216–221. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3573381.3596467 3, 4

[50] Y. Wei, R. Shi, A. U. Batmaz, Y. Li, M. Huang, R. Yang, and H.-N. Liang. Evaluating and modeling the effect of frame rate on steering performance in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–14, 2024. doi: 10.1109/TVCG.2024.3451491 2, 4, 6

[51] Y. Wei, R. Shi, D. Yu, Y. Wang, Y. Li, L. Yu, and H.-N. Liang. Predicting gaze-based target selection in augmented reality headsets based on eye and head endpoint distributions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, article no. 283, 14 pages. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544548.3581042 2

[52] Y. Wei, K. Xu, Y. Li, L. Yu, and H.-N. Liang. Exploring and modeling directional effects on steering behavior in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 30(11):7107–7117, 2024. doi: 10.1109/TVCG.2024.3456166 1, 6

[53] S. Yamanaka. Steering performance with error-accepting delays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, 9 pages, p. 1–9. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290605.3300800 5

[54] S. Yamanaka and H. Miyashita. Modeling the steering time difference between narrowing and widening tunnels. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, 11 pages, p. 1846–1856. Association for Computing Machinery, New York, NY, USA, 2016. doi: 10.1145/2858036.2858037 5

[55] S. Yamanaka and H. Miyashita. Modeling pen steering performance in a single constant-width curved path. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*, ISS '19, 12 pages, p. 65–76. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3343055.3359697 4, 6

[56] S. Yamanaka and W. Stuerzlinger. The effect of latency on movement time in path-steering. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, article no. 622, 19 pages. Association for Computing Machinery, New York, NY, USA, 2024. doi: 10.1145/3613904.3642316 2, 4

[57] S. Yamanaka, T. Takaku, Y. Funazaki, N. Seto, and S. Nakamura. Evaluating the Applicability of GUI-Based Steering Laws to VR Car Driving: A Case of Curved Constrained Paths. *Proceedings of the ACM on Human-Computer Interaction*, 7(ISS):430:93–430:113, Nov. 2023. doi: 10.1145/3626466 1, 2, 6

[58] D. Yu, X. Lu, R. Shi, H.-N. Liang, T. Dingler, E. Velloso, and J. Goncalves. Gaze-supported 3d object manipulation in virtual reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, article no. 734, 13 pages. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3411764.3445343 3

[59] S. Zhai, J. Accot, and R. Woltjer. Human action laws in electronic virtual worlds: An empirical study of path steering performance in vr. *Presence: Teleoperators and Virtual Environments*, 13(2):113–127, 04 2004. doi: 10.1162/1054746041382393 5

[60] S. Zhai, C. Morimoto, and S. Ihde. Manual and gaze input cascaded (magic) pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, 8 pages, p. 246–253. Association for Computing Machinery, New York, NY, USA, 1999. doi: 10.1145/302979.303053 2

[61] X. Zhang, X. Ren, and H. Zha. Improving eye cursor's stability for eye pointing tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, 10 pages, p. 525–534. Association for Computing Machinery, New York, NY, USA, 2008. doi: 10.1145/1357054.1357139 6

[62] X. Zhang, X. Ren, and H. Zha. Modeling dwell-based eye pointing target acquisition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, 10 pages, p. 2083–2092. Association for Computing Machinery, New York, NY, USA, 2010. doi: 10.1145/1753326.1753645 2, 3, 7