The Effect of Latency on Movement Time in Path-steering

Shota Yamanaka*
syamanak@lycorp.co.jp
Yahoo Japan Corporation
Tokyo, Japan

Wolfgang Stuerzlinger
w.s@sfu.ca
Simon Fraser University
Vancouver, Canada

ABSTRACT

In current graphical user interfaces, there exists a (typically unavoidable) end-to-end latency from each pointing-device movement to its corresponding cursor response on the screen, which is known to affect user performance in target selection, e.g., in terms of movement time (MT). Previous work also reported that a long latency increases MTs in path-steering tasks, but the quantitative relationship between latency and MT had not been previously investigated for path-steering. In this work, we derive models to predict MTs for path-steering and evaluate them with five tasks: goal crossing as a preliminary task for model derivation, linear-path steering, circular-path steering, narrowing-path steering, and steering with target pointing. The results show that the proposed models yielded an adjusted $R^2 > 0.94$, with lower AICs and smaller cross-validation $RMSE$s than the baseline models, enabling more accurate prediction of MTs.

CCS CONCEPTS

• **Human-centered computing** → **HCI theory, concepts and models**; **Pointing**; **Empirical studies in HCI**.

KEYWORDS

Human motor performance, operational time prediction, graphical user interface

1 INTRODUCTION

1.1 Background

The nature of interactive systems introduces unavoidable end-to-end latency (or *lag*, *delay*) from a user's input to its corresponding output on the screen, for example, due to hardware-sensing needs, data communication to the CPU, device driver overhead, the application's own processing time, graphics generation, display refresh,

---

and even network delays, if a remote desktop is used [18, 82]. Pointing is one of the most fundamental operations in graphical user interfaces (GUIs) and latency is known to negatively affect its performance, e.g., by increasing the movement time MT and error rate ER [22, 37]. Previous work has focused on the effects of latency on pointing [31, 56], and some of these studies have also derived predictive models how MT changes due to latency [37, 56].

Yet, users perform also operations other than pointing in GUIs, and the importance of path-steering tasks has been repeatedly identified [2, 7, 24, 44, 103, 109]. A steering task requires users to pass through a restricted path without the cursor deviating outside of the boundaries, e.g., when navigating a cascaded menu. The MT of this operation is well-modeled by the steering law, which is based on the path width and length [2, 26, 70].

In contrast to the numerous papers focusing on target pointing under the influence of latency, few studies have investigated the increase of MTs in steering tasks when the cursor movement is delayed. Furthermore, no theoretical model on MT has been presented, and thus it is currently necessary to conduct a new study for each individual latency to be able to estimate the MT; this is where a theoretical contribution of a performance model has its benefits.

In this paper, we propose modified formulations of the steering law to account for the presence of latency, and evaluate their validity in four experiments. As a necessary preliminary step to derive the modified steering law model, we also test a goal-crossing task with latency. Our contributions are twofold.

- Theoretical derivation of MT prediction models that take the end-to-end latency in path-steering tasks into account. Our basic model is derived from Fitts' law for constant-width paths with latency and then refined for each specific condition, such as narrowing paths and steering-and-pointing tasks.
- Empirical evaluation of the proposed models compared with baselines through four steering experiments and one goal crossing study. In all five experiments, the results showed that models that incorporate a linear latency term predict MTs significantly more accurately than all baseline models.

1.2 Contribution to the HCI Community and Suggestions to Future Researchers

Given that latencies in hardware and software are unavoidable, HCI researchers have tried to understand whether there is regularity to the effects of latency on user performance, whether these effects are predictable, and if predictive models can be utilized to prevent performance degradation. However, the HCI community's specific efforts in addressing these questions have been somewhat narrow and often limited to target-pointing tasks, despite repeated emphasis in previous studies on the importance of trajectory-based GUI

tasks such as goal crossing and path steering. This gap needs to be filled for better understanding of user performance in interactive systems.

Given that users regularly use many different devices today, the need for investigating the effects of latency on user performance is increasing. Typical HCI research on GUI-performance modeling has assumed that we use a single input-output device combination, where the end-to-end latency is on average constant when validating performance models, which is today an overly simplistic assumption. If we use several devices with different latencies simultaneously, e.g., in a dual-display environment or when using a remote desktop in parallel to the local one, the prediction accuracy of a model will be considerably degraded, and one could question whether existing model-based interaction techniques [46, 78] and model-driven design guidelines [13, 105] would still work effectively.

As our five experiments show below, realistic levels of latency significantly affect *MT*s. This result advances our understanding in HCI, by filling the above-mentioned gap on GUI performance; our models robustly explain the impact of latency on user behavior for steering tasks that are not modeled by Fitts' law. Beyond the scope of Fitts' and steering laws our takeaway for general HCI researchers is that there is still a need to re-evaluate the effects of latencies, even in well-known models. Examples include replicating model evaluations for *ER*s in pointing at stationary targets [97, 110], *MT*s in goal-crossing tasks [10, 54], and *MT*s in free-form object-selection tasks in illustration software [14, 101]. From this perspective, our work is a cornerstone for further studies that investigate models with latency for tasks different than pointing, to evaluate, e.g., if these models retain high prediction accuracy across multiple displays with different latencies or if incorporating a latency term into the model is necessary. At a more general level and when considering the plethora of devices that users use today, ranging from smartwatches to Virtual/Augmented reality systems, our work enables designers to predict how efficiently steering operations will work in a given system with a given latency.

2 RELATED WORK

In this section, we describe previous work on the effects of end-to-end latency for indirect input devices. We also mention studies on touchscreen-response latencies [42, 61, 93], robot-arm manipulation with latency [9, 15, 28, 36, 57, 74, 77], and delayed teleoperations [16, 69]. However, there are large differences between these works and the topic we are investigating. Thus, readers who are interested in robotics-related topics or touchscreens are directed to the above-mentioned papers.

2.1 Latency in Computer Operations

To identify the negative effects of latency on target pointing, Teather et al. used a motion capture system to measure the time from a mouse movement to its on-screen cursor displacement, and showed that *MT* and spatial error (i.e., distance between the cursor and target) increased with latency [82]. Pavlovych and Stuerzlinger showed that a longer latency affected *MT*, *ER*, and *throughput* (as defined for Fitts' law [76]) negatively [66].
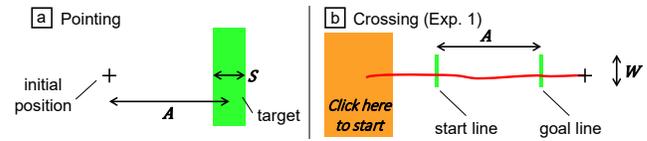
Figure 1: The standard formulation of Fitts' law holds for (a) target-pointing and (b) goal-crossing tasks, if there is no added latency.

In a series of studies, Claypool et al. reported similar negative effects of latency on moving-target pointing [20–22, 51]. Pavlovych et al. showed that in the pursuit of a moving target, a longer latency increased the spatial error [65, 68]. Long and Gutwin experimented with video games of Pong and Space Invaders, and found that latency had a negative impact on the user experience and performance, including increased *ER* and movement-path variability [52, 53].

Other work measured actual end-to-end latencies in computer systems. For example, Teather et al. reported that a Microsoft optical mouse had an end-to-end latency of 30 ± 2 ms [82]. Ivkovic et al. measured latencies in various video game titles, devices, and consoles such as WiiU [41]. The reported latencies ranged from 23 to 243 ms, and specifically were between 23 and 158ms for PC-and-mouse conditions. Casiez et al. reported that, using a red-LED mouse, the end-to-end latencies differed depending on the mouse, display, and OS, exhibiting between 50 and 83 ms of latency [18]. When users extend/mirror a secondary display via a wireless connection, latencies can range from 240 to 400 ms using the Microsoft Wireless Display Adapter [33], about 400 ms for Miracast [62], and approximately 120 ms for AirPlay [1].

2.2 Fitts' law and a Modified Version that Accounts for Latency

The *MT* to point to a target in tasks such as the ones shown in Figure 1a is well-modeled by Fitts' law [30, 55]:

$$MT = a + b \cdot ID_f \ \text{ and } \ ID_f = \log_2\left(\frac{A}{S} + 1\right), \tag{1}$$

where ID_f is the Fitts' index of difficulty (as indicated by the subscript), A is the distance to the target center, and S is the target size. In this paper, italic lower-case letters $a{-}c$ and those with prime (e.g., c') indicate empirically determined coefficients. Equation 1 also holds for a goal-crossing task (Figure 1b), where the user needs to to pass through a line, where it is sufficient to replace S with the line length W, i.e., $ID_f = \log_2\left(\frac{A}{W} + 1\right)$ [5, 10].

MacKenzie and Ware [56] and Hoffmann [37] independently derived mathematically equivalent models versions of Fitts' law by adding a latency term:

$$MT = a + b \cdot ID_f + c \cdot L_{\text{total}} \cdot ID_f = a + b \cdot (1 + c' \cdot L_{\text{total}}) \cdot ID_f, \tag{2}$$

where L_{total} is the total end-to-end latency that sums a system's inherent and (programmatically) added latencies, with $c' = c/b$. Their derivations were also similar; reaching a target consists of several discrete sub-movements, and the latency increases each sub-movement's time, resulting in the linear increase in ID_f and
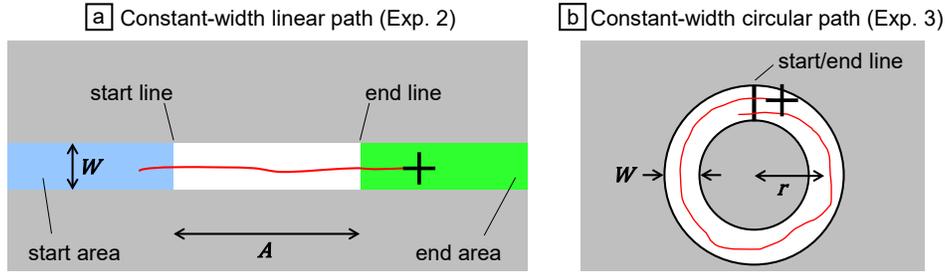
Constant-width linear path (Exp. 2)   Constant-width circular path (Exp. 3)

Figure 2: Without programmatically added latency, the steering law holds for (a) constant-width linear and (b) circular paths.

thus this model linearly increases MT. Typical Fitts' law experiments use a single apparatus and thus L_{total} is constant on average. Consequently, Equation 2 is equivalent to the baseline model of Equation 1.

MacKenzie and Ware used a mouse, examined L_{total} = 8.3, 25, 75, and 225 ms, and reported that the baseline model showed $R = 0.560$ while their proposed model showed $R = 0.967$. The smallest L_{total} = 8.3 ms was explained as "half the screen refresh period of 16.67 ms," but Pavlovych and Stuerzlinger identified that MacKenzie and Ware's experiment likely had a system latency of more than 60 ms [67]. Still, the model fit is not affected by such an additional constant latency because L_{total} is a linear term, while the coefficient values change accordingly. To identify this issue more clearly and assuming that the apparatus's inherent, unavoidable latency is "L_{base}", such as 50 ms [18], we call a latency added by the experimental program "L_{added}". The total latency is then the sum of both these term $L_{total} = L_{added} + L_{base}$.

Ware and Balakrishnan tested Equation 2 for pointing in a virtual reality environment with hand- and head-motion tracking systems, and reported $R^2 = 0.90$ [90]. Hoffmann used a potentiometric recorder, and the participants rotated a knob to move a pen onto a target on the monitor [37]. He tested L_{total} = 30, 200, 500, and 1000 ms, and reported that Equation 2 showed $R^2 = 0.97$. Hoffmann and Karri later conducted a follow-up study using a mouse with L_{added} = 0, 300, 600, 1000, and 2000 ms, and reported that Equation 2 showed $R^2 = 0.97$ after excluding the L_{added} = 0-ms data [39]. They also provided a reanalysis of So et al.'s study on head motion-based pointing with L_{added} = 0, 133, and 267 ms [75], obtaining $R^2 = 0.97$ by Equation 2 under reasonably difficult conditions.

Friston et al. examined very low-latency conditions using L_{added} = 0, 10, 20, 30, 50, and 80 ms, with L_{base} = 6 ms in their mouse-tracking system [31]. Equation 2 showed a poor fit of $R^2 = 0.384$, and thus they claimed that this model holds when the latency is sufficiently long to be noticeable for users. This matches observations by Pavlovych and Stuerzlinger, where a drop-off of performance was only observed above (about) 60 ms [67].

2.3 Steering Law and Path-steering Tasks with Latency

The steering law is a model to predict MT to pass through a path whose width is W and the length is A [2, 26, 70]:

$$MT = a + b \cdot ID_s \text{ and } ID_s = \frac{A}{W}, \qquad (3)$$

where ID_s is the index of difficulty for the steering law (indicated by corresponding subscript). This model applies to both linear and circular paths (see Figure 2, where $A = 2\pi r$ for the circular one) [4, 38, 112] and various types of input devices [3, 72].

Friston et al. tested a path-steering task with L_{added} ranging from 0 to 80 ms, and reported that MT and ER increased as L_{added} increased [31]. However, the paths were manually drawn and twisted substantially with parameters A and W not being controlled. Tochioka et al. conducted an experiment in which the participants dragged a circle on a touchscreen to pass through three straight-path segments connected with two corners [85]. The image of the user's hand and the screen was captured by an overhead camera, and the participants viewed it with latency (L_{added} ranging 4.3–154.3 ms). They reported that MT increased for $L_{added} \geq 64.3$ ms, but did not mention the steering-law fit.

The most recent work in the field by Wiese and Henze examined negative latencies [91]. This was achieved by using artificial neural networks to predict the position of the cursor in the future. They used straight paths and examined L_{total} values of -50, -16.67, 0, $+16.67$, and $+50$ ms, where smaller latencies decreased MT more. For each latency, the steering law held with $R^2 = 0.83$–0.93, but the model fit across all latency conditions was not reported (only for individual fits).

2.4 Steering Tasks under Other Types of Latency

Drury measured the pen-movement speed to steer through a circular path drawn on a piece of paper, with the light in the experiment room being repeatedly turned on/off [27]. To avoid deviating from the path, the participants could not move the pen during the dark period, and thus they had to stand by for the light being turned on, resulting in a delayed operation. Later, similar studies have been conducted using an indirect input pen tablet, where (instead of the room light turning on/off) the mouse cursor appeared/disappeared [48] or using a direct input pen tablet where the screen was turned on/off [47, 49]. In these studies, the light/cursor/display were blinking, but no additional end-to-end latency was added because the participants could see the current pen/cursor position in real time during the bright period.

Yamanaka et al. evaluated another type of latency, which allows the cursor to deviate from the path for a short period, such as 400 ms [35, 95]. However, in their experiments, the system responded to cursor movements and clicking operations immediately, and thus there is little relationship of their work with our current study.

2.5 Summary of Previous Results

Previous work consistently found that latency negatively affects quantitative performance for pointing such as MT and ER [21, 31, 53, 56] and subjective feelings [52]. Also, in moving-target tracking tasks, the deviation from the intended trajectory increases with latency [65, 68], which might suggest that users have to be more careful in path-steering tasks in long latency conditions to avoid deviating from path boundaries.

Even though we looked even at research on robot arms and teleoperations, some of which investigated Fitts' tasks [28, 36, 57, 74], we found no work on the accuracy of models of the effect of latency on the steering law, i.e., whether it applies to MT data regardless of latency. The most relevant previous studies are those of Friston et al. [31] and Tochioka et al. [85]. They investigated path-steering tasks with several levels of added latency L_{added}, but did not report the steering-law fits. The exception is the work by Wiese and Henze who reported R^2 values for each of their artificially generated L_{total} values from -50 to $+50$ ms [91]. A larger L_{total} increased MT from 610 to 680 ms, which motivated us to investigate the goodness of fit under more clearly noticeable latency conditions, i.e., exceeding $+60$ ms [67].

Drury [27] and Lin et al. [48, 49] investigated situations where the participants could not see the path or cursor for a given period. However, the relationship between their blinking conditions and our current interest in end-to-end latency is unclear or, at least, there is no evidence that their results are guaranteed to apply. In summary, while the path-steering performance under latency has been measured, there are no quantitative models to predict steering times, even though researchers have emphasized the necessity of modeling operational times with latency [31, 91].

3 EXPERIMENTS

3.1 Choice of Crowdsourced User Experiments

We conducted all five experiments on the *Yahoo! Crowdsourcing* platform [94]. The experimental systems were developed in the Hot Soup Processor programming language and set to run at 200 Hz (i.e., in a 5-ms loop). The crowdworkers downloaded and ran an executable file to perform a given task. Our affiliation's IRB-equivalent ethics team and the crowdsourcing platform approved this study[1].

Our choice of using a crowdsourcing service was motivated by the relative ease of recruitment of a large number of participants. Also, the steering law and our modified models are designed to predict average MTs exhibited by a group of participants. Thus, using a larger number of participants performing more trials will increase the likelihood that the models are more representative. On the other hand, one concern with crowdsourcing is that there is less control over the experimental environment (e.g., the device, lack of distractions in the room) and instruction compliance compared

to laboratory experiments. Yet, this approach also increases the external validity of the results, as it involves participants performing tasks in real environments.

This might introduce some issues for model evaluation; if a modified model provides the same level of prediction accuracy as the baseline model, this outcome might be due to the lack of control in the experiment rather than our model's inferiority. However, previous studies have reported that for GUI model validity crowd-sourced experiments lead to conclusions that are consistent with laboratory-based ones, e.g., for Fitts' law [29, 43, 97], a scrolling-time model [71], rectangular-target pointing models [96], or the steering law for linear and circular paths [107]. We see this as evidence that our hypothesis (that the proposed models fit the MT data more accurately than baseline formulations) can be tested accurately through crowdsourced experiments.

If a potential future laboratory-based replication study reaches a different conclusion than our work, this would provide evidence in the ongoing discussion of comparisons of laboratory vs. crowd-sourcing environments, e.g., whether a model fit changes when screen sizes are controlled or not, as previous work has investigated [29, 43]. The topic of laboratory vs. crowdsourcing experiments is worthy of investigation but is beyond the scope of our current study, which focuses on model derivation and validation in a crowd-sourced environment.

3.2 Participants and Recruitment

We recruited crowdworkers who used Windows Vista or later versions and a display having a resolution of at least 1280×720 pixels to run our experimental software. We used an option in the crowd-sourcing platform for screening newly created accounts, which enabled us to offer the task only to workers who were – on the basis of their previous task history – considered reliable (but without their task history details being disclosed).

To reduce the noise that would be introduced by using a variety of input devices, we asked workers to use a mouse if available. Nevertheless and to avoid potential biases, we still mentioned that any pointing device was acceptable but then simply removed data from non-mouse users from the analysis. The mouse specifications (e.g., DPI, wired/wireless) were not controlled. To increase ecological validity, we also did not require any changes to the cursor speed or pointer-acceleration function in the Control Panel. When displays with different pixel densities or mice with different DPIs are used, the distance in mm that the cursor moves on the screen varies for a fixed mouse movement. However, previous studies found that, after merging the data from numerous participants who used unknown displays, mice, and possibly with different pointer-acceleration functions, and then analyzing the average MT, both Fitts' and steering laws still yielded good fits [29, 97, 107]. Thus, we expect similar outcomes for our proposed models.

Once workers accepted the task, they were asked to read the text-based instructions and watch a 30-sec demo video in which one of the authors performed the task. After they had then finished all experimental trials and completed a follow-up questionnaire, they uploaded the log data file to a server to receive payment.

---

[1]According to the decision by the ethics team and legal department of our affiliation, data can be disclosed after statistical summarizing, e.g., by reporting mean and 95% CI of MT. In contrast, for example, the following results and demographics cannot be mentioned: "two workers were left-handed," "six workers used Windows 8," and "the task-completion times ranged from 10 min and 1 sec to 31 min and 7 sec" as these are related to individuals' results. This decision does not affect our research conclusion, since we only discuss the prediction accuracy of the average MT data, as in previous studies on the steering law.

4 EXPERIMENT 1: GOAL CROSSING TASK

The goal-crossing task is the basis for all our modified models on path-steering tasks. Because Fitts' law holds for crossing tasks [2, 5, 10, 54], one could simply assume that the version of Fitts' law that accounts for latency (Equation 2) also holds for a crossing task with latency. Yet, as there is no real evidence to support such a speculation in the literature, we decided to conduct an empirical test to strengthen our model derivation for steering tasks.

4.1 Participants

In total, 38 mouse-users completed this experiment. The reward was JPY 300 (~USD 2.28 as of April 2023). The average task duration was 16 min 35 sec, and thus the mean effective hourly payment was JPY 1,107 (~USD 8.42).

4.2 Task

The task window (1200×700 pixels) depicted two vertical target lines (start and goal, respectively) with length W lying at a distance A (Figure 1b). The task was to click within the orange start area located at the left edge of the window and then cross the two lines as rapidly and accurately as possible. Movement direction was always to the right. The crosshair cursor left a red trace. If the cursor missed the start line (i.e., passed over an extension of the start line), the participants had to click the start area again to retry the same target condition. This was not counted as a valid trial, as we consider that the core goal-crossing task had not been begun at this stage. We (also) forced participants who passed through the start line and then missed the goal line, i.e., performed an erroneous trial, to redo the same trial. After crossing the goal line, the results for each trial (the time elapsed and the number of retrials) were shown in the start area. The next pair of target lines appeared whenever the user moved the cursor back to the left half of the window.

4.3 Design and Procedure

Experiment 1 used a $5 \times 2 \times 3$ within-subjects design: five L_{added} values (0, 50, 100, 150, and 200 ms), two As (420 and 700 pixels), and three Ws (11, 23, and 51 pixels). The ID_f ranged from 3.21 to 6.01 bits, which covered a range from easy to difficult task conditions [76]. Because we could not measure L_{base} for each worker's apparatus, on the basis of [18] we used $L_{\text{base}} = 50$ ms as an approximation, resulting in L_{total} ranging from 50 to 250 ms. We deliberately set L_{total} to adequately cover a range of typical local latencies of mouse-and-PC systems, such as 50–83 ms [18], 28–32 ms [82], and 23–158 ms [41], but our L_{total} also covers latencies typical for network-connected displays, e.g., 240–400-ms [33].

We measured two dependent variables: MT and ER. MT was the time between crossing the start and goal lines. Because failed trials were redone from the beginning, and erroneous trials were ignored, the MT data involved only error-free trials. We defined the ER as the percentage of the number of missed goal lines over the number of total valid trials.

For a fixed L_{added} condition, the participants repeatedly performed seven repetitions for each of the $2_A \times 3_W$ conditions that appeared in a random order; 42 trials in total. The first repetition was considered as practice, and the remaining six repetitions (i.e., 36 trials) were used for data collection. Thus, the participants performed

180 data-collection trials ($36_{\text{trials}} \times 5_{L_{\text{added}}}$). Before a new L_{added} condition began, a large circular button informed participants of the change. The order of the five L_{added} values was randomized. In total, we analyzed the data from $5_{L_{\text{added}}} \times 2_A \times 3_W \times 6_{\text{repetitions}} \times 38_{\text{participants}} = 6,840$ trials.

4.4 Result of Experiment 1

While we followed a conventional approach to analyze the data statistically, such as running ANOVAs and pairwise tests, we only list the results that relate to our main research question (model-fit comparison) in this section. Other, secondary results, including F-values, p-values, and η_p^2 values, are listed in the supplementary materials. Overall, our results were consistent with previous studies such as that MT increased with the task difficulty and L_{total}, and thus the participants performed the task appropriately as intended in this experiment.

4.4.1 Data Screening to Remove Outliers. We used the inter-quartile range (*IQR*) method to detect trial-level outliers and removed trials with extremely fast or slow operations, as done in previous work [29]. The *IQR* is defined as the difference between the first and third quartiles of the MT for each of the task conditions ($5_{L_{\text{added}}} \times 2_A \times 3_W$) for each participant. Trials in which the MT deviated by more than $3 \cdot IQR$ from the first or third quartile [25] were removed. Among the 6,840 error-free trials, we identified 124 trial-level outliers (1.81%).[2]

To identify participant-level outliers, we calculated the mean MT across all task conditions for each participant. We applied the *IQR* method using each participant's mean MT to identify extremely rapid or slow participants, but found no participant-level outliers.
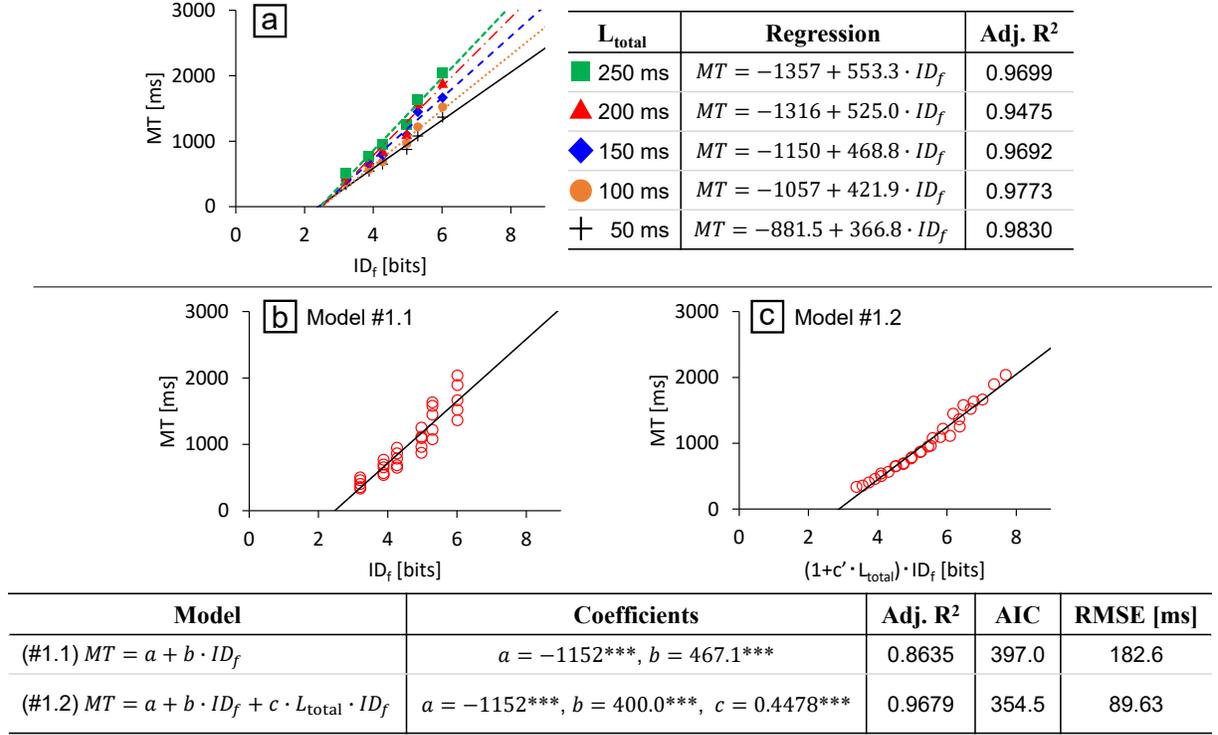
4.4.2 Model Fit. We first applied the baseline Fitts' law to the six corresponding (= $2_A \times 3_W$) fitting points for each L_{total} condition. Figure 3a shows that Fitts' law held with adjusted $R^2 > 0.94$, which is close to results for pointing ($R^2 > 0.93$ [37, 75]). The fact that the results show negative intercepts is also consistent with previous work on crossing tasks [10, 54, 86, 92, 102].

Next, we compared the two candidate models using all 30 (= $5_{L_{\text{total}}} \times 2_A \times 3_W$) fitting points. One is the baseline Fitts' law model (Equation 1 denoted as Model #1.1 in Figure 3b), and the other is the version with the added latency term (Equation 2 denoted as Model #1.2 in Figure 3c). As mentioned above, we used $L_{\text{base}} = 50$ ms, but note that this choice did not affect the model fit.

Because Models #1.1 and #1.2 have different numbers of coefficients, we use an adjusted R^2 measure (higher is better). In addition, to compare the model fit statistically, we computed the AIC (smaller is better) [8]. As a brief guideline, (a) a model with $AIC \leq (AIC_{\text{minimum}} + 2)$ is probably comparable with a better model; and (b) a model with $AIC \geq (AIC_{\text{minimum}} + 10)$ should be rejected [17].

Model #1.2 yielded a higher adjusted R^2 and a significantly smaller AIC than Model #1.1. Also, Model #1.2 showed significance $p < 0.001$ for all coefficients. To create the scatter plot shown in

---

[2]When the number of repetitions per task condition is small, outlier detection is not that robust, and the *IQR* method may identify a large number of outliers, whereas other criteria may identify fewer outliers. To check this, we applied the 3σ method, which excluded fewer outliers, but still confirmed that our main conclusion holds: models with the latency term can predict MT more accurately. See Section 6 of the Supplementary Material for details.

L_{total}	Regression	Adj. R^2
■ 250 ms	$MT = -1357 + 553.3 \cdot ID_f$	0.9699
▲ 200 ms	$MT = -1316 + 525.0 \cdot ID_f$	0.9475
◆ 150 ms	$MT = -1150 + 468.8 \cdot ID_f$	0.9692
● 100 ms	$MT = -1057 + 421.9 \cdot ID_f$	0.9773
＋ 50 ms	$MT = -881.5 + 366.8 \cdot ID_f$	0.9830

Model	Coefficients	Adj. R^2	AIC	RMSE [ms]
(#1.1) $MT = a + b \cdot ID_f$	$a = -1152^{***}, b = 467.1^{***}$	0.8635	397.0	182.6
(#1.2) $MT = a + b \cdot ID_f + c \cdot L_{total} \cdot ID_f$	$a = -1152^{***}, b = 400.0^{***}, c = 0.4478^{***}$	0.9679	354.5	89.63

Figure 3: Model-fit results of Experiment 1. (a) Fitts' law regressions for each L_{total}. Results using all 30 fitting points for (b) the baseline version of Fitts' law and (c) the model with an added latency term. Throughout this paper, the p-values for coefficients are annotated as "$*$" $p < 0.05$, "$**$" $p < 0.01$, and "$***$" $p < 0.001$.

Model	Adj. R^2	AIC	RMSE [ms]
(#1.a) $MT = a + b \cdot L_{total}$	0.0566	455.0	477.2
(#1.b) $MT = a + b \cdot ID_f + c \cdot L_{total}$	0.9541	365.2	109.0
(#1.c) $MT = a + b \cdot L_{total} + c \cdot L_{total} \cdot ID_f$	0.8918	391.0	168.5

Table 1: Model-fit result of additional candidate formulations in Experiment 1.

Figure 3c, we transformed the regression expression as follows:

$$MT = -1152 + 400.0 \cdot ID_f + 0.4478 \cdot ID_f \cdot L_{total}$$
$$= -1152 + 400.0 \cdot (1 + 0.0011195 \cdot L_{total}) \cdot ID_f. \quad (4)$$

With this transformation and when visually comparing Figure 3b with Figure 3c, we can confirm that more data points are located closer to the regression line.

Lastly, to evaluate the prediction accuracy for a new (untested) task condition, we ran a leave-one-condition-out cross-validation (LOOCV). For this cross-validation, we predicted the MT of one fitting point by using the coefficients computed from a regression for the remaining 29 points, and repeated this process for all the 30 points. Then, we computed the root mean square error $RMSE$ across all 30 observed and predicted MTs. Model #1.1 yielded $RMSE = 182.6$ ms, while Model #1.2 gave 89.63 ms, indicating the better prediction accuracy for #1.2.

We conclude that introducing the latency term yields a significantly better model fit due to higher adjusted R^2, smaller AIC,

significant p-values for coefficients, smaller $RMSE$ of LOOCV, and less data-point deviation from the regression line, as visible in the corresponding graph. This result supports that the same latency term as for target-pointing tasks also yields significantly better prediction accuracy for a goal-crossing task, strengthening our model derivation for path-steering tasks described in the next section.

4.4.3 Comparison with Other Potential Model Formulations. While we do not have theoretical derivations for them, we tested three additional model variants to provide a better understanding of the influence of latency on MT. The first new candidate has only a latency term:

$$\text{Candidate model (a)} : MT = a + b \cdot L_{total} \quad (5)$$

The second one does not use the interaction term of $L_{total} \cdot ID_f$, but models latency independently:

$$\text{Candidate model (b)} : MT = a + b \cdot ID_f + c \cdot L_{total} \quad (6)$$
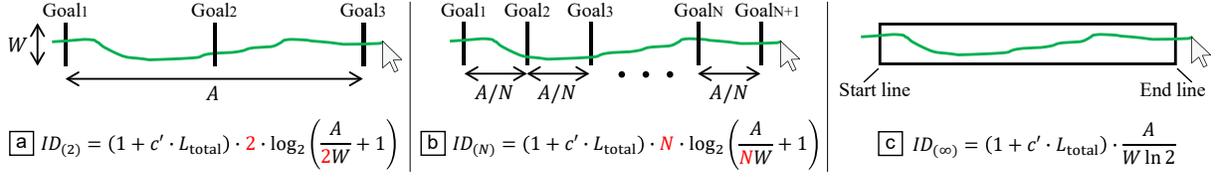
Figure 4: Model derivation for steering through a linear path with latency.

The third and last model does not include just the ID_f term but also the L_{total} term and the interaction term $L_{\text{total}} \cdot ID_f$:

$$\text{Candidate model (c)} : MT = a + b \cdot L_{\text{total}} + c \cdot L_{\text{total}} \cdot ID_f \quad (7)$$

As in the main analysis, we set $L_{\text{base}} = 50$ ms. Table 1 shows that each model-fit metric (adjusted R^2, AIC, and $RMSE$) was worse than those for our proposed Model #1.2. Thus, we empirically confirmed that there is no justification for the three new candidates.

5 EXPERIMENT 2: STEERING THROUGH CONSTANT-WIDTH LINEAR PATHS

Here, we first derive a model for the MT to steer through a constant-width linear path in the same manner as Accot and Zhai's development of the steering law [2] and then empirically evaluate the model.

5.1 Model Derivation

According to the result of Experiment 1, the time of a single crossing motion $MT_{(1)}$ to cross two lines with latency is

$$MT_{(1)} = a + b \cdot ID_{(1)}, \text{ where } ID_{(1)} = (1 + c' \cdot L_{\text{total}}) \cdot \log_2\left(\frac{A}{W} + 1\right). \quad (8)$$

We then put another goal line in the center of the two original lines (Figure 4a). When a user successively crosses Goal_1, Goal_2, and Goal_3, the movement distance is $A/2$ for each of the two goal-crossing tasks, and thus the total difficulty of the task $ID_{(2)}$ is

$$\begin{aligned} ID_{(2)} &= 2 \times \left[(1 + c' \cdot L_{\text{total}}) \cdot \log_2\left(\frac{A}{2W} + 1\right) \right] \\ &= (1 + c' \cdot L_{\text{total}}) \cdot 2 \cdot \log_2\left(\frac{A}{2W} + 1\right). \end{aligned} \quad (9)$$

In the same manner, when there are $N + 1$ lines (Figure 4b), the movement distance for each part is A/N, and the total task difficulty $ID_{(N)}$ is

$$ID_{(N)} = (1 + c' \cdot L_{\text{total}}) \cdot N \cdot \log_2\left(\frac{A}{NW} + 1\right). \quad (10)$$

When $N \to \infty$, the task turns into steering through a constrained path (Figure 4c). Using a first-order Taylor series expansion of $\log_2(x + 1)$ yields

$$ID_{(\infty)} = (1 + c' \cdot L_{\text{total}}) \cdot \frac{A}{W \ln 2}. \quad (11)$$

Accordingly, the MT for steering through a path with a delayed cursor is $MT = a + b \cdot ID_{(\infty)}$. Because $\ln 2$ is a constant that can be absorbed into the slope, we have

$$MT = a + b \cdot (1 + c' \cdot L_{\text{total}}) \cdot \frac{A}{W} = a + b \cdot \frac{A}{W} + c \cdot L_{\text{total}} \cdot \frac{A}{W}, \quad (12)$$

where $c' = c/b$. Since A/W is the steering-law difficulty ID_s, we can reformulate this to

$$MT = a + b \cdot (1 + c' \cdot L_{\text{total}}) \cdot ID_s = a + b \cdot ID_s + c \cdot L_{\text{total}} \cdot ID_s. \quad (13)$$

As a result, our model states that L_{total} linearly relates to ID_s, which is the same as the latency model of Fitts' law.

5.2 Participants, Task, Design, and Procedure

In total, 38 mouse-users completed this experiment. We did not restrict duplicate participation in all five experiments, and thus it is possible that the same worker participated more than once. Each of them received a reward of JPY 300. With an average task duration of 18 min 42 sec, the mean effective hourly rate was JPY 963 (~USD 7.32).

The task was to click on the blue start area, move the cursor along the white path, and then click on the green end area (Figure 2a). Movement direction was always to the right. If the cursor deviated from the path, participants had to try the same task condition again, starting from re-clicking in the start area. After each successful trial, a large circular button labeled "Next", which also displayed the results for the current trial (the time and number of failed attempts), appeared at the bottom-left of the window, and participants needed to click on it to reveal the next path condition. Participants were asked to complete each trial as quickly and accurately as possible.

Experiment 2 had a $5 \times 2 \times 3$ within-subjects design: five L_{added} values (0, 50, 100, 150, and 200 ms), two As (480 and 640 pixels), and three Ws (19, 29, and 47 pixels). ID_s ranged from 10.2 to 33.7, corresponding to a range of tasks that are challenging enough to require continuous visually controlled movements [72, 84].

For a fixed L_{added} condition, the participants performed seven repetitions for each $2_A \times 3_W$ conditions, which appeared in random order. The first repetition was again considered to be practice. In total, we analyzed the data from $5_{L_{\text{added}}} \times 2_A \times 3_W \times 6_{\text{repetitions}} \times 38_{\text{participants}} = 6{,}840$ trials.

5.3 Result of Experiment 2

Among the 6,840 error-free trials the IQR method identified 135 trial-level outliers (1.97%), but we found no participant-level outliers. Figure 5a shows that the baseline steering law held well for each L_{total} condition with adjusted $R^2 > 0.98$. Comparing the two candidate models using all 30 ($= 5_{L_{\text{total}}} \times 2_A \times 3_W$) fitting points, we found that according to the AIC measure the steering law (Model #2.1 in Figure 5b) yielded a significantly worse fit than the version with the latency term (Model #2.2 in Figure 5c).

Similar to the result in Experiment 1, the MT difference increased with ID_s due to the effect of L_{total} (Figure 5a). While almost all data points are located on or very close to the regression line for the

L_{total}	Regression	Adj. R^2
🟩 250 ms	$MT = -226.3 + 90.09 \cdot ID_s$	0.9889
🔺 200 ms	$MT = -293.3 + 89.34 \cdot ID_s$	0.9913
🔷 150 ms	$MT = -224.1 + 70.45 \cdot ID_s$	0.9930
🟠 100 ms	$MT = -256.0 + 68.59 \cdot ID_s$	0.9916
+ 50 ms	$MT = -215.8 + 61.23 \cdot ID_s$	0.9936

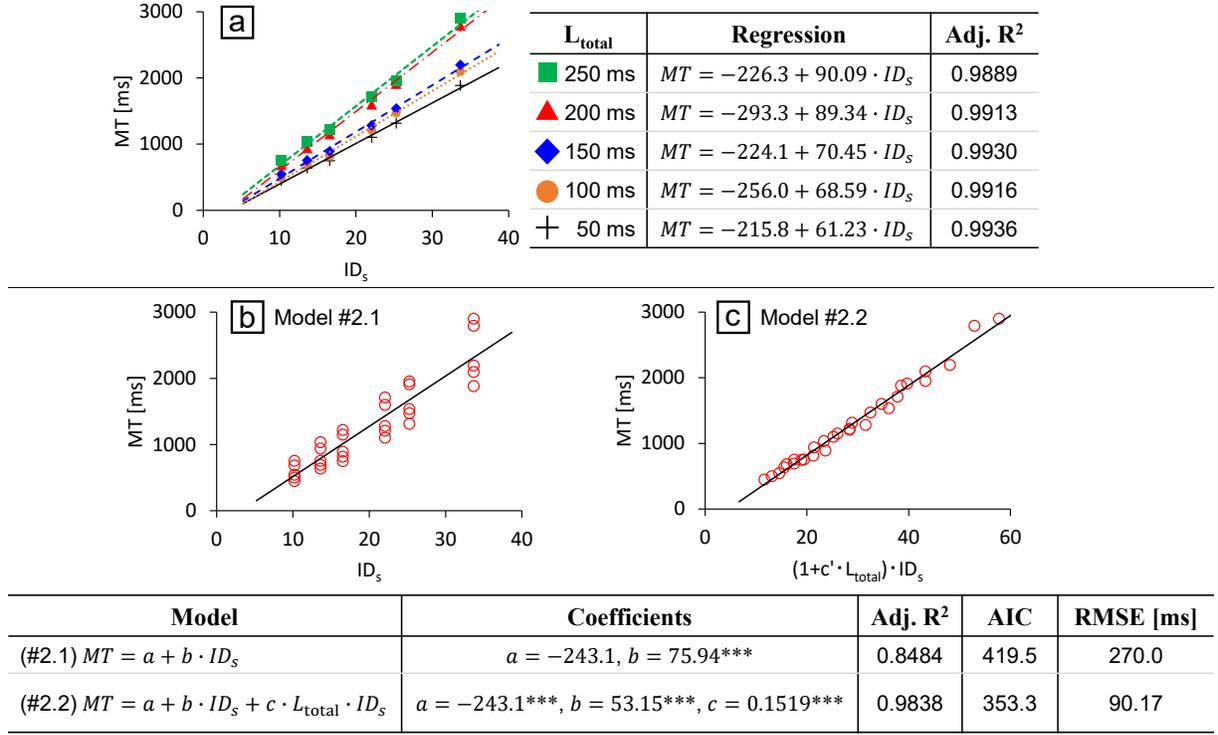| Model	Coefficients	Adj. R^2	AIC	RMSE [ms]
(#2.1) $MT = a + b \cdot ID_s$	$a = -243.1,\ b = 75.94^{***}$	0.8484	419.5	270.0
(#2.2) $MT = a + b \cdot ID_s + c \cdot L_{\text{total}} \cdot ID_s$	$a = -243.1^{***},\ b = 53.15^{***},\ c = 0.1519^{***}$	0.9838	353.3	90.17

Figure 5: Model-fit results of Experiment 2. (a) Steering law regressions for each L_{total}. Results using all 30 fitting points for (b) the baseline and (c) latency models.

Model	Adj. R^2	AIC	RMSE [ms]
(#2.a) $MT = a + b \cdot L_{\text{total}}$	0.0814	473.5	650.1
(#2.b) $MT = a + b \cdot ID_s + c \cdot L_{\text{total}}$	0.9643	377.0	136.5
(#2.c) $MT = a + b \cdot L_{\text{total}} + c \cdot L_{\text{total}} \cdot ID_s$	0.9045	406.5	218.3

Table 2: Model-fit result of additional candidate formulations in Experiment 2.

model with the latency term, as shown in Figure 5c, this phenomenon could not be captured by the baseline steering law (Figure 5b). Model #2.2 exhibited $p < 0.001$ for all coefficients, and the three prediction-accuracy metrics (adjusted R^2, AIC, and RMSE) all support that our new model better predicted MTs than the baseline steering law.

Table 2 lists the model fits for the three additional candidate formulations, which are derived in the same way as in Experiment 1 (Section 4.4.3). The results were worse than those for our proposed Model #2.2, and again did not support the three new candidates.

6 EXPERIMENT 3: STEERING THROUGH CONSTANT-WIDTH CIRCULAR PATHS

Steering through a circular path requires dynamical changes in the movement direction, and thus requires a significantly larger MT and ER than a linear path under the same ID_s condition [3, 4, 79, 111]. Previous work has shown that the baseline steering law model applies to linear and circular paths, both theoretically [38, 58] and empirically [3, 79, 111]. Thus it is logical to assume that the modified

model with latency (Equation 13) can model the MTs in tasks to steer through circular paths better than the baseline, and we verify this assumption in this experiment.

6.1 Participants, Task, Design, and Procedure

In total, 36 mouse-users completed this experiment. For this experiment, the reward was JPY 300. The task duration was 17 min 28 sec on average, and thus the mean effective hourly payment was JPY 1,031 (~USD 7.83).

The task was to click on any point to the left of the top start line, move the cursor along the white path clockwise, cross the top end line, and then click somewhere on the right side of the end line (Figure 2b). Since steering through a circular path takes up to three time as long as a linear path under the same ID_s condition [3, 111], we had to reduce the numbers of independent-variable levels and repetitions to adapt the task to our crowdsourced experiment. As it has been well-demonstrated that the steering law holds for circular paths, we decided not to explore a variety of ID_s levels at this stage.

L_{total}		Regression	Adj. R^2
■	250 ms	$MT = -115.6 + 397.6 \cdot ID_s$	0.9908
▲	200 ms	$MT = +134.7 + 334.2 \cdot ID_s$	0.9935
◆	150 ms	$MT = -244.2 + 316.3 \cdot ID_s$	0.9907
●	100 ms	$MT = -203.6 + 260.9 \cdot ID_s$	0.9905
+	50 ms	$MT = -298.2 + 252.0 \cdot ID_s$	0.9932

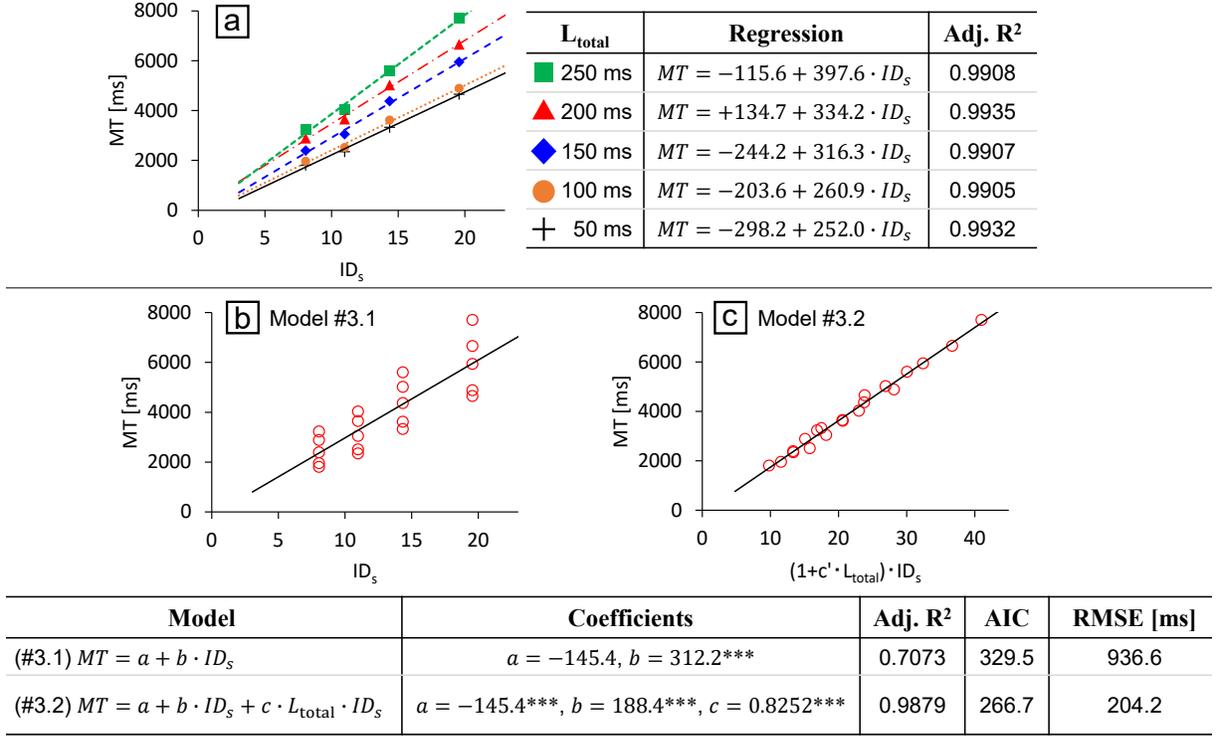| Model	Coefficients	Adj. R^2	AIC	RMSE [ms]
(#3.1) $MT = a + b \cdot ID_s$	$a = -145.4, b = 312.2^{***}$	0.7073	329.5	936.6
(#3.2) $MT = a + b \cdot ID_s + c \cdot L_{total} \cdot ID_s$	$a = -145.4^{***}, b = 188.4^{***}, c = 0.8252^{***}$	0.9879	266.7	204.2

Figure 6: Model-fit results of Experiment 3. (a) Steering law regressions for each L_{total}. Results using all 20 fitting points for (b) the baseline and (c) the latency models.

Model	Adj. R^2	AIC	RMSE [ms]
(#3.a) $MT = a + b \cdot L_{total}$	0.2053	349.5	1526
(#3.b) $MT = a + b \cdot ID_s + c \cdot L_{total}$	0.9663	287.1	343.9
(#3.c) $MT = a + b \cdot L_{total} + c \cdot L_{total} \cdot ID_s$	0.9263	302.8	488.8

Table 3: Model-fit result of additional candidate formulations in Experiment 3.

In comparison to Experiment 2, we thus used only two shorter As and two wider Ws, resulting in four smaller ID_s values.

Experiment 3 used a $5 \times 2 \times 2$ within-subjects design: five L_{added} values (0, 50, 100, 150, and 200 ms), two As (330 and 450 pixels), and three Ws (23 and 41 pixels). ID_s ranged from 8.05 to 19.6. The number of repetitions was six, including the first one, which we considered as a practice trial. In total, we analyzed the data from $5_{L_{added}} \times 2_A \times 2_W \times 5_{repetitions} \times 36_{participants} = 3,600$ trials.

6.2 Result of Experiment 3

Among the 3,600 error-free trials, the IQR method identified 136 trial-level outliers (3.78%). We found no participant-level outliers. We observed an adjusted $R^2 > 0.99$ for each L_{total} using the baseline steering law (Figure 6a). When all 20 conditions (= $5_{L_{total}} \times 2_A \times 2_W$) were analyzed simultaneously, the baseline steering law (Model #3.1 in Figure 6b) exhibited a significantly worse fit than the latency model (Model #3.2 in Figure 6c) judging on the basis of the adjusted R^2, AIC, and $RMSE$ measures, see Figure 6. Also, Model #3.2 yielded

$p < 0.001$ for all coefficients, and thus the latency term is necessary to predict MTs more accurately than the baseline steering law.

Table 3 lists the model fits for the three additional candidate formulations. The results were worse than those for our proposed Model #3.2, and again document that the three new candidates are not supported.

7 EXPERIMENT 4: STEERING THROUGH NARROWING PATHS

Steering tasks sometimes require dynamical changes in speed, e.g., when the current path width changes, which occurs when users need to select multiple objects through a single stroke, i.e., via lasso selection [106]. We thus develop a model for narrowing paths on the basis of Accot and Zhai's derivation [2].

7.1 Model Derivation

As shown in Figure 7, we define the widths of the start and end lines as W_{start} and W_{end}, respectively, where $W_{start} > W_{end}$. According to Accot and Zhai [2], the index of difficulty of a narrowing path
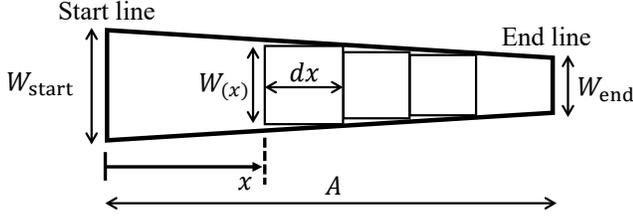
Figure 7: Model derivation for steering through a narrowing path.

ID_n is

$$ID_n = \frac{A}{W_{\text{end}} - W_{\text{start}}} \ln \frac{W_{\text{end}}}{W_{\text{start}}}. \tag{14}$$

We revisit the derivation of this model to develop a modified formulation to account for latency. Dividing a narrowing path into infinitesimal sections whose length is dx, we consider the difficulty in steering through one section at the position x. The width of this section is $W(x) = W_{\text{start}} + x(W_{\text{end}} - W_{\text{start}})/A$. Based on the result of our Experiment 2, the task difficulty with latency dID_x of this section can then be modeled as the movement distance divided by the width, and thus we obtain

$$dID_x = (1 + c' \cdot L_{\text{total}}) \cdot ID_s$$
$$= (1 + c' \cdot L_{\text{total}}) \cdot \frac{dx}{W_{\text{start}} + \frac{x}{A}(W_{\text{end}} - W_{\text{start}})}. \tag{15}$$

To compute the MT to steer through the entire narrowing path, we integrate dID_x along the path for $0 \le x \le A$:

$$MT = a + b \int_0^A (1 + c' \cdot L_{\text{total}}) \cdot \frac{dx}{W_{\text{start}} + \frac{x}{A}(W_{\text{end}} - W_{\text{start}})}$$
$$= a + b \cdot (1 + c' \cdot L_{\text{total}}) \cdot \frac{A}{W_{\text{end}} - W_{\text{start}}} \ln \frac{W_{\text{end}}}{W_{\text{start}}}. \tag{16}$$

The resulting derived formulation is similar to the latency model for constant-width paths (Equation 13):

$$MT = a + b \cdot (1 + c' \cdot L_{\text{total}}) \cdot ID_n = a + b \cdot ID_n + c \cdot L_{\text{total}} \cdot ID_n. \tag{17}$$

7.2 Participants, Task, Design, and Procedure

In total, 32 mouse-users completed this experiment. They received JPY 300 as compensation. With an average task duration of 16 min 15 sec the mean effective hourly payment was JPY 1,108 (~USD 8.42).

The task, design, and procedure were almost identical to Experiment 2. The exception was that we used three W_{start} values and one W_{end} of 16 pixels. This experimental design, with a fixed W_{end} matches Accot and Zhai's work [2].

Experiment 4 used a $5 \times 2 \times 3$ within-subjects design: five L_{added} values (0, 50, 100, 150, and 200 ms), two As (480 and 640 pixels), and three W_{start} values (32, 64, and 96 pixels). ID_n ranged from 10.8 to 27.7. The first of the six repetitions was again counted as practice, leaving five data points for analysis. In total, we analyzed the data from $5_{L_{\text{added}}} \times 2_A \times 3_{W_{\text{start}}} \times 5_{\text{repetitions}} \times 32_{\text{participants}} = 4{,}800$ trials.

7.3 Result of Experiment 4

Among the 4,800 error-free trials, the IQR method identified 166 trial-level outliers (3.46%). We found no participant-level outliers. The baseline ID_n model held with adjusted $R^2 > 0.92$ for each L_{total} (Figure 8a). When all 30 conditions (= $5_{L_{\text{total}}} \times 2_A \times 3_{W_{\text{start}}}$) were analyzed simultaneously, the baseline Model #4.1 (Figure 8b) resulted in a significantly worse fit than the latency model (Model #4.2 in Figure 8c), on the basis of adjusted R^2, AIC, and $RMSE$ measures. Also, Model #4.2 yielded $p < 0.001$ for all coefficients, and thus we can state that the inclusion of the latency term enabled us to predict MTs more accurately than the baseline ID_n model.

Table 4 lists the model fits for the three additional candidate formulations. The results were worse than those for our proposed Model #4.2, and again do not support the three new candidates.

8 EXPERIMENT 5: TARGETED-STEERING TASK

Experiments 2–4 used a single path segment, but in actual GUIs, other operations may require a combination of actions to be performed. For example, in cascaded-menu navigation, after passing through the desired parent menu item, users then can click on one of its submenu items. Such an operation is called a targeted-steering task [45, 73] and several models that combine Fitts' and steering law terms have been proposed. In Experiment 5, we derive several candidate models for targeted-steering with latency and evaluate them.

8.1 Model Derivation

8.1.1 Baseline Models for Targeted-steering Task. In this task, as shown in Figure 9a, users click in the start area, pass through the white path, and then click on the green target. We measure the time from entering the white path to clicking on the target. Kulikov and Stuerzlinger's model is the simplest candidate, where the task difficulty is the sum of those for steering and pointing:

$$MT = a + b \cdot ID_{\text{sum}},$$
$$\text{where } ID_{\text{sum}} = ID_s + ID_f = \frac{A}{W} + \log_2\left(\frac{A + S/2}{S} + 1\right). \tag{18}$$

The second candidate is Dennerlein et al.'s model, which assumes that the gradients of ID_s and ID_f for MT are unequal:

$$MT = a + b_s \cdot ID_s + b_f \cdot ID_f, \tag{19}$$

where b_s and b_f are the slopes on the steering and Fitts' law regressions, respectively. Both models define the target distance of Fitts' law as $A + S/2$, which is the total length of the path and the distance to the target center.

According to Thibbotuwawa et al., for steering through a linear path, when the path is sufficiently short or wide ($A < n \times W$, where $n \approx 5$), the steering motion can be completed with a ballistic motion using no visual feedback [84]. Based on this, Senanayake et al. hypothesized that once the cursor reaches $5W$ before the path end, users do not need to pay attention to the path boundaries for the remainder of the path, and can concentrate on the pointing motion [73]. Thus, the distance for the steering-law difficulty is
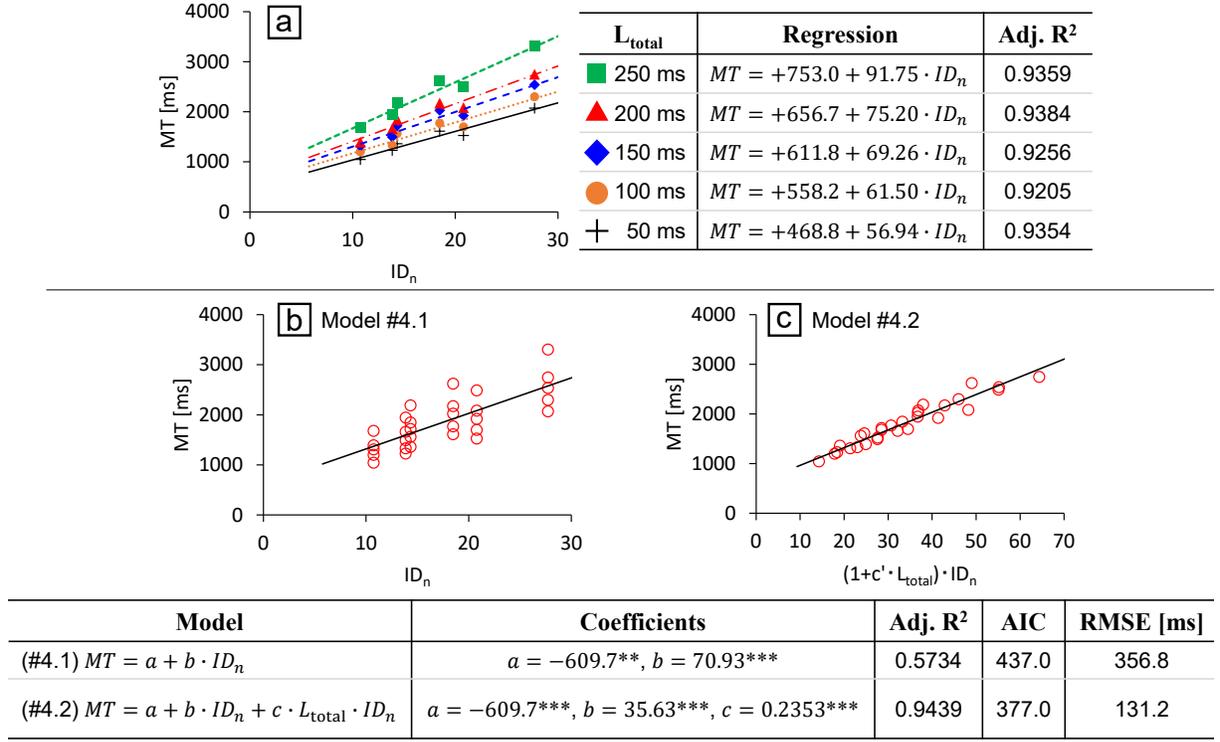
L_{total}	Regression	Adj. R^2
🟩 250 ms	$MT = +753.0 + 91.75 \cdot ID_n$	0.9359
🔺 200 ms	$MT = +656.7 + 75.20 \cdot ID_n$	0.9384
🔷 150 ms	$MT = +611.8 + 69.26 \cdot ID_n$	0.9256
🟠 100 ms	$MT = +558.2 + 61.50 \cdot ID_n$	0.9205
+ 50 ms	$MT = +468.8 + 56.94 \cdot ID_n$	0.9354

Model	Coefficients	Adj. R^2	AIC	RMSE [ms]
(#4.1) $MT = a + b \cdot ID_n$	$a = -609.7^{**}, b = 70.93^{***}$	0.5734	437.0	356.8
(#4.2) $MT = a + b \cdot ID_n + c \cdot L_{total} \cdot ID_n$	$a = -609.7^{***}, b = 35.63^{***}, c = 0.2353^{***}$	0.9439	377.0	131.2

Figure 8: Model-fit results of Experiment 4. (a) Steering law regressions for each L_{total}. Results using all 30 fitting points for (b) the baseline and (c) the latency models.

Model	Adj. R^2	AIC	RMSE [ms]
(#4.a) $MT = a + b \cdot L_{total}$	0.3232	450.9	446.8
(#4.b) $MT = a + b \cdot ID_n + c \cdot L_{total}$	0.9298	383.8	148.5
(#4.c) $MT = a + b \cdot L_{total} + c \cdot L_{total} \cdot ID_n$	0.8990	394.7	175.6

Table 4: Model-fit result of additional candidate formulations in Experiment 4.
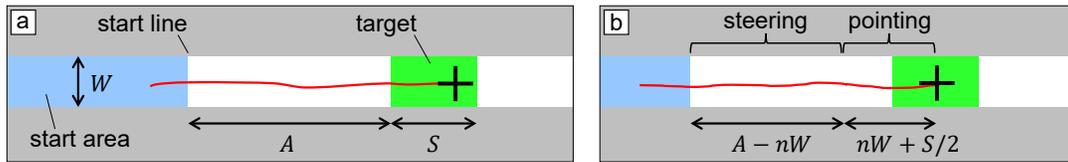

Figure 9: (a) Parameters in a targeted-steering task. (b) Senanayake's model assumes that a steering motion ends nW pixels before entering the target.

$A - nW$, and the distance for Fitts' law is $nW + S/2$ (see Figure 9b):

$$MT = a + b_s \cdot \frac{A - nW}{W} + b_f \cdot \log_2 \left(\frac{nW + S/2}{S} + 1 \right) \quad (20)$$

This model yielded $R^2 = 0.869$ when both the steering and pointing tasks require visually controlled movements.

8.1.2 Modified Models in the Presence of Latency. According to the result of our Experiment 2 and previous studies on Fitts' law with latency, the latency term has linear effects on both difficulties in the steering and Fitts' laws. Thus, introducing the effect of latency into Kulikov and Stuerzlinger's model yields the following formulation.

$$MT = a + b \cdot (1 + c' \cdot L_{total}) \cdot ID_{sum} = a + b \cdot ID_{sum} + c \cdot L_{total} \cdot ID_{sum}, \quad (21)$$

It is unclear whether the effects of L_{total} are the same for ID_s and ID_f, thus the modified version of the model of Dennerlein et al. is

$$MT = a + b_s \cdot (1 + c'_s \cdot L_{total}) \cdot ID_s + b_f \cdot (1 + c'_f \cdot L_{total}) \cdot ID_f$$
$$= a + b_s \cdot ID_s + c_s \cdot L_{total} \cdot ID_s + b_f \cdot ID_f + c_f \cdot L_{total} \cdot ID_f, \quad (22)$$

where c_s is a coefficient for the interaction term of the latency and ID_s, c_f is that of the latency and ID_f, $c'_s = c_s/b_s$, and $c'_f = c_f/b_f$. Lastly, the modified version of Senanayake et al. model is

$$
\begin{aligned}
MT = a + b_s \cdot (1 + c'_s \cdot L_{\text{total}}) \cdot \frac{A - nW}{W} \\
+ b_f \cdot (1 + c'_f \cdot L_{\text{total}}) \cdot \log_2\left(\frac{nW + S/2}{S} + 1\right) \\
= a + b_s \cdot \frac{A - nW}{W} + c_s \cdot L_{\text{total}} \cdot \frac{A - nW}{W} \\
+ b_f \cdot \log_2\left(\frac{nW + S/2}{S} + 1\right) + c_f \cdot L_{\text{total}} \cdot \log_2\left(\frac{nW + S/2}{S} + 1\right).
\end{aligned}
\tag{23}
$$

In our model fitting analysis, we test both a fixed value of $n = 5$ and an optimized n. Since $n = 5$ was fixed based on a previous study, Senanayake et al. considered it not to be a free parameter, i.e., $n = 5$ does not change depending on the observed MT values to maximize R^2. We followed this when we calculated adjusted R^2 and AIC, while we considered it is a free parameter when optimizing n.

8.2 Participants, Task, Design, and Procedure

In total, 35 mouse-users completed this experiment. Each received an amount of JPY 400 for the task, which lasted on average 25 min 48 sec. Thus the mean effective hourly rate was JPY 930 (~USD 7.07).

Experiment 5 utilized a $4 \times 2 \times 2 \times 3$ within-subjects design: four L_{added} values (0, 67, 133, and 200 ms), two As (480 and 640 pixels), two Ws (19 and 47 pixels), and three Ss (15, 25, and 45 pixels). Because we used the additional independent variable S, we reduced the number of levels for L_{added}. $ID_s = A/W$ ranged from 10.2 to 33.7, and $ID_f = \log_2[(A + S/2)/S + 1]$ ranged from 3.60 to 5.46 bits. For a fixed L_{added}, the participants performed six repetitions (including a first one for practice) for the $2_A \times 2_W \times 3_S$ conditions that appeared in a random order. The order of the four L_{added} values was also randomized. In total, we analyzed the data from $4_{L_{\text{added}}} \times 2_A \times 2_W \times 3_S \times 5_{\text{repetitions}} \times 35_{\text{participants}} = 8,400$ trials.

8.3 Result of Experiment 5

Among the 8,400 error-free trials, the IQR method identified 299 trial-level outliers (3.56%). We found no participant-level outliers.

8.3.1 Model Fit for each L_{total} Condition.
Figure 10 shows the fits for each model for each L_{total}. Kulikov's model showed comparatively lower adjusted R^2 values than the other three candidates. In contrast, Dennerlein's and Senanayake's models exhibited > 0.99 under all L_{total} conditions. For the regression graphs of Dennerlein's and Senanayake's models, the ID values on the horizontal axis are obtained by dividing the second and subsequent terms by the second coefficient. For example, under the $L_{\text{total}} = 50$ ms condition of Dennerlein's model, we have

$$
\begin{aligned}
MT &= -527.3 + 49.79 \cdot ID_s + 254.0 \cdot ID_f \\
&= -527.3 + 49.79 \cdot (ID_s + 5.101 \cdot ID_f),
\end{aligned}
\tag{24}
$$

where the term inside the parentheses is shown on the horizontal axis as the "ID" value.

In Senanayake's model with an optimized n, for $L_{\text{total}} = 50, 117, 183$, and 250 ms, we obtained $n = 6.727, 37.95, 10.18$, and 15.27,

respectively, with adjusted $R^2 > 0.99$ for all. Yet, the results of $n = 37.95$ or 15.27 are inappropriate with respect to the structure of Senanayake's model; if $n = 37.95$ is correct, then the pointing motion would have started 1784 pixels before the path end under the $W = 47$-pixel condition, which is practically infeasible as A was up to 640 pixels and the experimental system window's width was only 1200 pixels.

Therefore, we decided to impose the constraint that $nW \leq A$ under all task conditions. With this constraint, we obtained $n = 10.21$ for $L_{\text{total}} = 50$ and 183 ms. This gives the upper limit for n when W is 47 pixels, and thus the steering difficulty was zero. Figure 10d shows the fits with this constraint.

Accordingly, the use of n as an extra free parameter led to the interpretation that the participants did not perform steering motions for several task conditions, although we had explicitly set A/W to always be greater than 10, requiring a visually controlled steering motion. Overall, we did not confirm a clear benefit of using an extra free parameter in Senanayake's model, since a fixed $n = 5$ from the previous study already exhibited a similar prediction accuracy to when using an optimized n.

8.3.2 Model Fit for all L_{total} Conditions.
Figure 11a–d shows the model fits for each model without the latency term (Equations 18–20), and Figure 11e–h with the latency terms (Equations 21–23). First, we visually confirmed that the data points are further away from the regression lines for the models without the latency term (Models #8.1–8.4). In contrast, for Models #8.5–8.8, the data points are close to the line and thus the MTs could be predicted more accurately. The better prediction accuracy for Models #8.5–8.8 is also validated through the measures of adjusted R^2, AIC, and $RMSE$.

We need to be careful when interpreting Model #8.7 (fixed $n = 5$) in comparison to #8.8 (optimized n). Given that the optimal value of n was 7.419, the reason why Model #8.7 yielded the best adjusted R^2 and AIC was due to the fact that a fixed $n = 5$ is somewhat close to 7.419, while $n = 5$ was not counted as a free parameter, as explained earlier. We are currently unable to determine from this experiment alone whether Model #8.7 would work better than #8.8 in other studies or application contexts.

We then attempted to identify the best model among all candidates, particularly Models #8.6, #8.7, and #8.8. The difference in adjusted R^2 for these models was trivial (less than 0.0005). The difference in $RMSE$ values was less than 3 ms (less than one frame), and thus we cannot argue for the superiority of any model. Finally, the AIC difference was less than 2 and we thus could not confirm a statistically significant difference in prediction accuracy. From these results, it is logical to conclude that we cannot make a clear decision which one the best model is. On the other hand, when comparing Models #8.5–8.8 with #8.1–8.4, all of the prediction-accuracy metrics were improved. This indicated that our proposed method worked well, i.e., modifying the existing models by integrating L_{total} to linearly increase the ID values led significant improvements in predicting MT.

We also examine the new three candidates for each of the four models: Kulikov and Stuerzlinger, Dennerlein et al., Senanayake et al. ($n = 5$), and Senanayake et al. (optimized n). For Senanayake et al.'s model, we name its steering-difficulty part $ID_{SenaS} = \frac{A - nW}{W}$
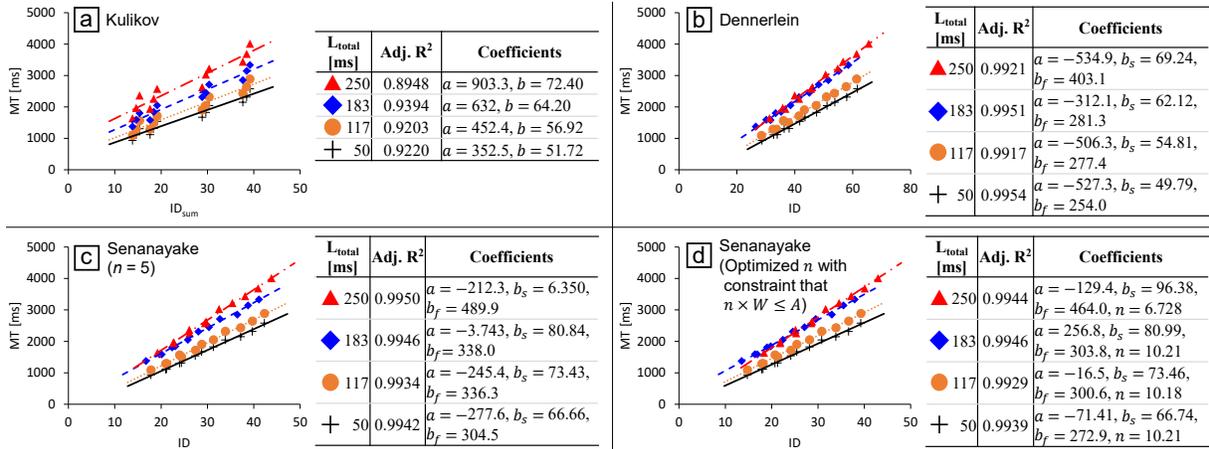
Figure 10: Fitting results for each candidate model in Experiment 5.

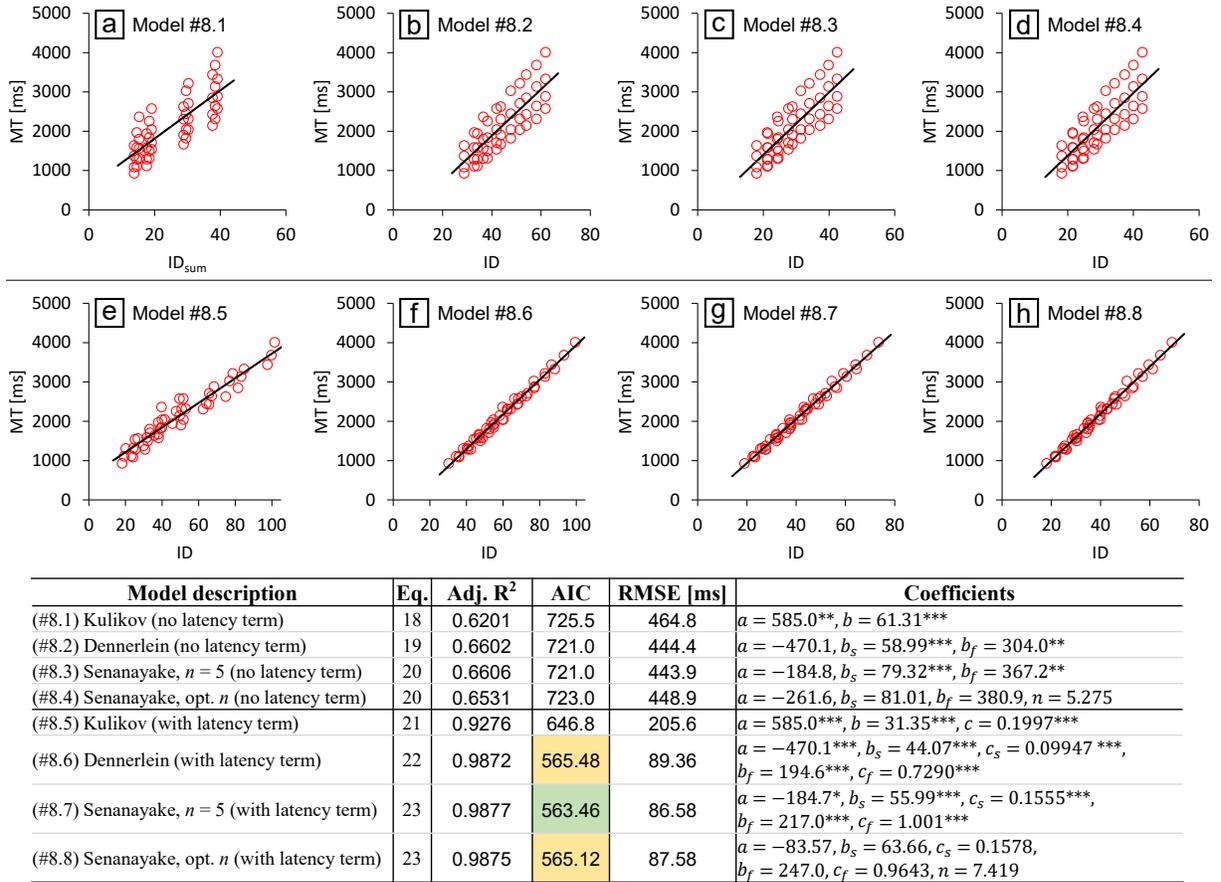| Model description	Eq.	Adj. R^2	AIC	RMSE [ms]	Coefficients
(#8.1) Kulikov (no latency term)	18	0.6201	725.5	464.8	$a = 585.0{**}, b = 61.31{***}$
(#8.2) Dennerlein (no latency term)	19	0.6602	721.0	444.4	$a = -470.1, b_s = 58.99{***}, b_f = 304.0{**}$
(#8.3) Senanayake, $n = 5$ (no latency term)	20	0.6606	721.0	443.9	$a = -184.8, b_s = 79.32{***}, b_f = 367.2{**}$
(#8.4) Senanayake, opt. n (no latency term)	20	0.6531	723.0	448.9	$a = -261.6, b_s = 81.01, b_f = 380.9, n = 5.275$
(#8.5) Kulikov (with latency term)	21	0.9276	646.8	205.6	$a = 585.0{***}, b = 31.35{***}, c = 0.1997{***}$
(#8.6) Dennerlein (with latency term)	22	0.9872	565.48	89.36	$a = -470.1{***}, b_s = 44.07{***}, c_s = 0.09947{***},$ $b_f = 194.6{***}, c_f = 0.7290{***}$
(#8.7) Senanayake, $n = 5$ (with latency term)	23	0.9877	563.46	86.58	$a = -184.7{*}, b_s = 55.99{***}, c_s = 0.1555{***},$ $b_f = 217.0{***}, c_f = 1.001{***}$
(#8.8) Senanayake, opt. n (with latency term)	23	0.9875	565.12	87.58	$a = -83.57, b_s = 63.66, c_s = 0.1578,$ $b_f = 247.0, c_f = 0.9643, n = 7.419$

Figure 11: Model-fit results of Experiment 5. The p-values cannot be calculated for the nonlinear regressions used in Senanayake's model with optimized n. The green cell shows the best result, and yellow ones are close-to-best results according to AIC.

and the Fitts-difficulty part $ID_{SenaF} = \log_2\left(\frac{nW+S/2}{S} + 1\right)$. Thus, the baseline model without the latency term can be written as

$$MT = a + b_s \cdot ID_{SenaS} + b_f \cdot ID_{SenaF}. \qquad (25)$$

We then modified this model in the same manner as those for Experiments 1–4. The results are listed in Tables 5–8. The model-fit

Model	Adj. R^2	AIC	RMSE [ms]
(#5.a) $MT = a + b \cdot L_{\text{total}}$	0.2870	755.7	635.7
(#5.b) $MT = a + b \cdot ID_{\text{sum}} + c \cdot L_{\text{total}}$	0.9272	647.1	206.7
(#5.c) $MT = a + b \cdot L_{\text{total}} + c \cdot L_{\text{total}} \cdot ID_{\text{sum}}$	0.8657	676.5	280.1

Table 5: Model-fit result for additional candidate formulations of Kulikov and Stuerzlinger's model in Experiment 5.

Model	Adj. R^2	AIC	RMSE [ms]
(#5.d) $MT = a + b \cdot L_{\text{total}}$	0.2870	755.7	635.7
(#5.e) $MT = a + b_s \cdot ID_s + b_f \cdot ID_f + c \cdot L_{\text{total}}$	0.9752	596.2	122.9
(#5.f) $MT = a + b \cdot L_{\text{total}} + c_s \cdot L_{\text{total}} \cdot ID_s + c_f \cdot L_{\text{total}} \cdot ID_f$	0.9097	658.3	229.7

Table 6: Model-fit result for additional candidate formulations of Dennerlein et al.'s model in Experiment 5.

Model	Adj. R^2	AIC	RMSE [ms]
(#5.g) $MT = a + b \cdot L_{\text{total}}$	0.2870	755.7	635.7
(#5.h) $MT = a + b_s \cdot ID_{SenaS} + b_f \cdot ID_{SenaF} + c \cdot L_{\text{total}}$	0.9757	595.3	121.6
(#5.i) $MT = a + b \cdot L_{\text{total}} + c_s \cdot L_{\text{total}} \cdot ID_{SenaS} + c_f \cdot L_{\text{total}} \cdot ID_{SenaF}$	0.9104	658.0	228.8

Table 7: Model-fit result for additional candidate formulations of Senanayake et al.'s model with fixed value of $n = 5$ in Experiment 5.

Model	Adj. R^2	AIC	RMSE [ms]
(#5.g) $MT = a + b \cdot L_{\text{total}}$	0.2870	755.7	635.7
(#5.h) $MT = a + b_s \cdot ID_{SenaS} + b_f \cdot ID_{SenaF} + c \cdot L_{\text{total}}$	0.9753	597.0	122.4
(#5.i) $MT = a + b \cdot L_{\text{total}} + c_s \cdot L_{\text{total}} \cdot ID_{SenaS} + c_f \cdot L_{\text{total}} \cdot ID_{SenaF}$	0.8087	695.3	326.7

Table 8: Model-fit result for additional candidate formulations of Senanayake et al.'s model with the optimized value of n in Experiment 5.

measures showed that these additional formulations did not yield better prediction accuracy than those shown in Figure 11.

9 DISCUSSION

9.1 Effects of Added Latency on User Performance

In all five experiments, L_{added} had significant main effects on MT and ER. The effect sizes for MT ranged from $\eta_p^2 = 0.312$ to 0.778 (see the supplementary materials), which all correspond to *large effect sizes* [88], indicating solid results. When we analyzed the MT data across all L_{total} conditions, the models including the L_{total} term showed substantially better fits according to the results of higher adjusted R^2s, significantly lower AICs, better visual fits, and smaller $RMSE$s in cross-validation.

In all modified models in Experiments 1 through 5, L_{total} linearly increased the task difficulty. As a linear term does not affect the model fit, changing L_{base} from 50 ms only changes the coefficients, but does not affect the model-fit. Thus, the choice of L_{base} does not affect our conclusion that the modified models will always predict MTs more accurately than the baselines, regardless of the actual base latency of the system used.

Smaller $RMSE$ values exhibited by the modified models indicate that MTs can be more accurately predicted than the baselines, even for untested path parameters and L_{added} values, as long as they are within the investigated range (i.e., extrapolation is not guaranteed). For example, our proposed models will still be able to predict MT accurately for an untested condition with $L_{\text{added}} = 300$ ms, but experimental evidence is necessary to confirm this assumption.

We designed Experiments 2 to 5 with the goal of validating as many different elements of steering tasks as possible, but previous studies have examined additional factors, e.g., the direction of movement [83, 112], curvature [58, 60, 100], paths with corners [64, 87, 101], widening paths [98, 99], and conditions where multiple paths are connected [103, 104]. Even under these conditions, we expect MT to increase with L_{total}, but verification of further refined models requires additional experiments. It might be possible that any of these conditions could show counterexamples relative to our current conclusion, such as that L_{total} does not significantly affect MT. Such a new finding does not negate the results of our five experiments; rather, it would provide an additional contribution to deepen our understanding of steering performance with latency. Still, based on the evidence of our outcomes we believe this to be an unlikely outcome.

9.2 Implications

Given the lower frequency of steering tasks compared to pointing in interactive systems, we acknowledge that implications based on our models are likely less broad compared to those of Fitts' law.

However, as steering tasks are more challenging than pointing (exhibiting longer MTs and higher ERs), good steering-facilitation techniques are potentially strong contributions, and previous researchers have proposed several techniques based on the steering law.

For example, *Attribute gates* can activate multiple commands by performing a single stroke through them and the MT of this operation is modeled by the steering law [78]. *Sloppy selection* intelligently determines the selected objects when a user circles several objects in a note-taking tool, while estimating the imaginary path width that the user is trying to pass through [46]. More specifically, with lower movement speed, the width of the path boundaries formed by the objects that the user does not want to select are assumed to be closer, and thus the path narrower. However, as we have shown, the baseline steering law model does not predict the MT (thus movement speed) well for conditions with added latency.

Suppose that a user is operating a PC with a dual-display setup and that the two displays have different latencies, e.g., because the secondary display is connected through a network-based display adapter [33] in presentation scenarios, with a 240-ms latency. Depending on which screen the user is then working on, the objects selected by *Sloppy selection* would be different because the steering speed naturally changes due to the latency. If existing interaction methods could use a refined version of the steering law that takes latency into account, with appropriate model coefficients to match each user's environment, usability could be improved.

In a more common example and again in systems with unequal latencies on different displays, an effective design choice would be to improve usability by shortening the path length or increasing the path width within a GUI to keep the MT constant (or decrease it). For example, using the data from Experiment 2 ($MT = -243.1 + 53.15 \cdot ID_s + 0.1519 \cdot L_{total} \cdot ID_s$), we predict $MT = 769$ ms when $A = 500$, $W = 30$ pixels, and $L_{total} = 50$ ms in the primary display. However, if L_{total} is 240 ms, $MT = 1250$ ms is estimated for the same path. If $MT = 769$ ms is to be maintained on the secondary display, A must either be shortened to 339 pixels, W widened to 45 pixels, or any combination that yields the same ID_s. We developed a proof-of-concept application of this idea; see the supplementary materials.

For target pointing, a similar method to maintain MTs is currently already included in Microsoft Office. On the basis of Fitts' law item sizes in the Ribbon Menu are automatically adapted so that the pointing time is constant regardless of the display resolution [34]. This insight can also be implemented for path-steering tasks, for example, by obtaining the latency of the connected display from its device information (or through manual measurement) and appropriately widening the path width so that, for example, the navigation time in a cascaded menu remains constant.

Simulation results for this idea are shown in Figure 12a for the case where W varies. By using the data from Experiment 2 and (without loss of generality) fixing A to 500 pixels, we change W and obtain graphs for four L_{total} values. When W is large, the MTs are (relatively) close to each other regardless of the L_{total} value because the task is easy. However, when W is small, MT tends to increase with L_{total} because errors occur unless the speed is considerably reduced. Figure 12b shows this non-linear relationship in a 3D surface graph. Although such MT prediction might require
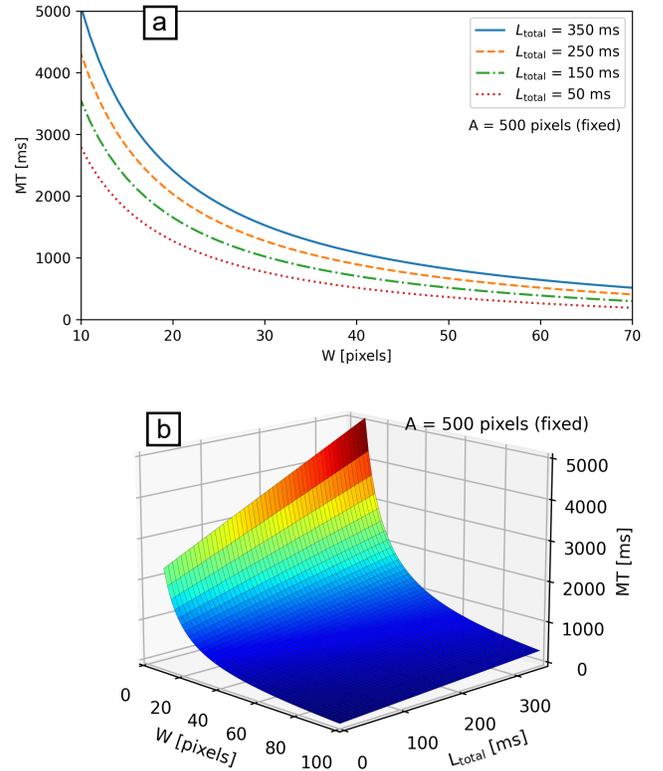

Figure 12: Simulations to predict MT by changing W and L_{total}.

a user test every time the path parameters and L_{total} are changed substantially, unless a quantitative model is available, our proposed models can reduce that cost and time.

9.3 Potential Applications for Future GUIs beyond Mouse-based Operations

We conducted all experiments only with mouse-based operations. However, the steering law is applicable to various situations and devices beyond the classical desktop environment. For example, in tasks where participants wear augmented reality (AR) glasses and perform a linear-path steering task with their head motions to support communication among doctors in surgery, the baseline steering law model held [12]. Another example is to use stereoscopic glasses and a motion-capturing stylus to steer through a 3D tunnel, for which a steering law with a curvature term holds [50]. This result is useful for predicting the difficulty of using 3D drawing tools in virtual reality (VR) spaces (e.g., [11, 19, 108]). For example, the task difficulty of the lasso selection tool in 2D drawing tools can be modeled by the steering law [87, 101], and thus it is reasonable to expect that the operational difficulty and time for drawing tools in 3D spaces could potentially be similarly predicted.

In immersive environments such as AR and VR, where operations are conducted using head or 3D stylus motions, it is valuable to verify whether our proposed models yield high prediction accuracy. This is relevant, as the end-to-end latency in VR environments is

today (generally) comparable to those of recent mouse and LCD display combinations. For example, in environments using the Unity game engine and SteamVR, the average latency with controllers, such as the HTC Vive, Oculus Rift, Oculus Rift S, and Valve Index ones, can range from 21 to 42 ms, but individually observed latencies can be even longer [82, 89], especially in AR systems, where latencies are typically higher [59]. In addition, network conditions could influence usability; e.g., in the above-mentioned steering tasks with head motions, the data transmission between HoloLens and Unity can happen either over USB or Wi-Fi [12]. If the Wi-Fi connection is not fast and stable, significant latency may occur and possibly increases the task difficulty.

As we believe that our proposed models are applicable to AR/VR tasks, the difficulty adjustment application that shortens the path's A or widens its W (Section 9.2) are likely directly relevant. For scenarios where changing the path's appearance is undesirable, expanding only the collision detection of boundaries in either spatial [80, 81] or temporal dimensions [35, 95] are potential alternatives. However, these techniques have only been validated through user experiments in 2D desktop environments, so the future development of 3D user interfaces requires the following further studies:

- Verifying whether our latency-inclusive models hold for AR/VR systems with head motions or handheld controllers.
- Conducting 3D versions in AR/VR environments for experiments that have been validated only in 2D desktop settings, such as the reduction in difficulty by widening the path [35, 80] or the time prediction of lasso selection tasks using the steering law [87, 101], is also needed.

9.4 Limitations

Our experiments used only a mouse as input device, but it is known that the steering and Fitts' laws hold for a variety of devices. We thus expect that our proposed models will also fit the data when other input devices are used. In addition, the specifications of the mice, PCs, and displays naturally varied across crowdworkers. Conducting follow-up experiments in a more controlled laboratory-based environment, using a single apparatus is thus a potential avenue for future work.

Our findings are also limited by the task parameters we tested. For latency, previous studies have examined values ranging from much shorter to substantially longer than those in our experiments, such as 6 ms [31] and 2,000 ms [39]. In contrast, our L_{total}s ranged from 50 to 250 ms, which covers the range from a typical mouse-and-display condition [18] to a short-ranged network-connected display extension [33]. Consequently, future work could explore up to what upper limit of latency our proposed models work effectively.

The steering and Fitts' laws are known to be applicable to a wide range of task difficulty levels, but if the task is too easy (when A is too short, W is too wide, or S is too large), the task can be completed with only a ballistic movement and MT depends then only on $\sqrt{A}$ [32, 84]. Also, for the rectangular target used in Experiment 5, previous studies have recommended using a linear [23] or nonlinear addition [6] depending on the target height. However, we limited our comparison to existing targeted-steering models based on the standard formulation of Fitts' law, because our research focus was to test the effects of latency on MT and model fit.

Lastly, our models are all empirical, grounded in experimental psychology and ergonomics. In contrast, recent development in machine learning (ML) have led to HCI studies using more complicated models that focus just on precise outcome prediction. For example, for gaze-based target selection, a model incorporating numerous parameters such as gaze coordinates and pupil diameter (127 in total) has been proposed, but it is acknowledged that discussing whether each coefficient is appropriate is infeasible [40]. In contrast, there are ML-based studies that engage in discussions about the appropriateness of each coefficient (cf. explainable ML). Park et al. proposed a reinforcement-learning model incorporating numerous parameters for pointing tasks, and they stated that "each parameter of the model has a clear cognitive meaning," and confirmed that each coefficient aligns with results from conventional cognitive psychology [63]. We also discussed in Section 8.3.1 that the value of $n = 37.95$ in Senanayake et al.'s model is inappropriate, which is possible because our tested models are interpretable. To enable future researchers who would like to focus on the prediction accuracy to reanalyze our data, the supplementary materials include the MT data for each task condition obtained in our experiments.

10 CONCLUSION

An end-to-end latency is a fundamental property of any interactive system, but the impact of such latency has only been examined for a limited set of interaction paradigms, mainly for target pointing. Yet, since various path-steering operations are also commonly used in GUIs, we investigated models that predict how MT increases with latency through four path-steering and one goal-crossing experiment. We derived models that assume that the end-to-end latency linearly increases the steering-law difficulty, and showed that MT could be predicted significantly more accurately than the baseline models in all cases. We hope that our work will help designers to predict the performance of interaction methods and GUI design in systems that are subject to notable latency.

REFERENCES

[1] Channel AAPL. 2021. AirPlay to Mac over Thunderbolt 3/Wi-Fi6 latency test. https://www.youtube.com/watch?v=inZWQ22NYug
[2] Johnny Accot and Shumin Zhai. 1997. Beyond Fitts' law: models for trajectory-based HCI tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '97)*. ACM, New York, NY, USA, 295–302. https://doi.org/10.1145/258549.258760
[3] Johnny Accot and Shumin Zhai. 1999. Performance Evaluation of Input Devices in Trajectory-based Tasks: An Application of the Steering Law. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Pittsburgh, Pennsylvania, USA) *(CHI '99)*. ACM, New York, NY, USA, 466–472. https://doi.org/10.1145/302979.303133
[4] Johnny Accot and Shumin Zhai. 2001. Scale Effects in Steering Law Tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, USA) *(CHI '01)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/365024.365027
[5] Johnny Accot and Shumin Zhai. 2002. More than Dotting the i's — Foundations for Crossing-Based Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Minneapolis, Minnesota, USA) *(CHI '02)*. Association for Computing Machinery, New York, NY, USA, 73–80. https://doi.org/10.1145/503376.503390
[6] Johnny Accot and Shumin Zhai. 2003. Refining Fitts' Law Models for Bivariate Pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI '03)*. ACM, New York, NY, USA, 193–200. https://doi.org/10.1145/642611.642646
[7] David Ahlström. 2005. 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* (Portland,

Oregon, USA) *(CHI '05)*. ACM, New York, NY, USA, 61–70. https://doi.org/10.1145/1054972.1054982

[8] Hirotugu Akaike. 1974. A new look at the statistical model identification. *IEEE Trans. Automat. Control* 19, 6 (Dec 1974), 716–723. https://doi.org/10.1109/TAC.1974.1100705

[9] Ryota Akimoto, Masatoshi Miyaji, Kenji Funahashi, Koji Tanida, and Shinji Mizuno. 2021. Positive Effect of Slight Delay for Operational Performance. In *10th IEEE Global Conference on Consumer Electronics, GCCE 2021, Kyoto, Japan, October 12-15, 2021*. IEEE, Washington, DC, USA, 162–166. https://doi.org/10.1109/GCCE53005.2021.9621805

[10] Georg Apitz, François Guimbretière, and Shumin Zhai. 2008. Foundations for Designing and Evaluating User Interfaces Based on the Crossing Paradigm. *ACM Trans. Comput.-Hum. Interact.* 17, 2, Article 9 (May 2008), 42 pages. https://doi.org/10.1145/1746259.1746263

[11] Rahul Arora, Rubaiat Habib Kazi, Fraser Anderson, Tovi Grossman, Karan Singh, and George Fitzmaurice. 2017. Experimental Evaluation of Sketching on Surfaces in VR. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5643–5654. https://doi.org/10.1145/3025453.3025474

[12] Takafumi Asao, Takeru Kobayashi, Kentaro Kotani, Satoshi Suzuki, Kazutaka Obama, Atsuhiko Sumii, and Tatsuto Nishigori. 2021. Development of a Drawing Application for Communication Support during Endoscopic Surgery and Its Evaluation Using Steering Law. *Applied Sciences* 11, 10 (2021), 13 pages. https://doi.org/10.3390/app11104505

[13] Xiaojun Bi and Shumin Zhai. 2016. Predicting Finger-Touch Accuracy Based on the Dual Gaussian Distribution Model. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. ACM, New York, NY, USA, 313–319. https://doi.org/10.1145/2984511.2984546

[14] Per Bjerre, Allan Christensen, Andreas K. Pedersen, Simon A. Pedersen, Wolfgang Stuerzlinger, and Rasmus Stenholt. 2017. Predictive Model for Group Selection Performance on Touch Devices. In *Human-Computer Interaction. Interaction Contexts*, Masaaki Kurosu (Ed.). Springer International Publishing, Cham, 142–161.

[15] James Hampton Black. 1971. Factorial study of remote manipulation with transmission of time delay. , 77 pages. Master's thesis.

[16] Claire Blackett, Alexandra Fernandes, Espen Teigen, and Thomas Thoresen. 2022. Effects of Signal Latency on Human Performance in Teleoperations. In *Human Interaction, Emerging Technologies and Future Systems V*, Tareq Ahram and Redha Taiar (Eds.). Springer International Publishing, Cham, 386–393.

[17] Kenneth P Burnham and David R Anderson. 2003. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media, Heidelberg, Germany.

[18] Géry Casiez, Stéphane Conversy, Matthieu Falce, Stéphane Huot, and Nicolas Roussel. 2015. Looking Through the Eye of the Mouse: A Simple Method for Measuring End-to-end Latency Using an Optical Mouse. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology* (Charlotte, NC, USA) *(UIST '15)*. ACM, New York, NY, USA, 629–636. https://doi.org/10.1145/2807442.2807454

[19] Noemie Chaniaud, Sylvain Fleury, Benjamin Poussard, Olivier Christmann, Thibaut Guitter, and Simon Richir. 2023. Is virtual reality so user-friendly for non-designers in early design activities? Comparing skills needed to traditional sketching versus virtual reality sketching. *Design Science* 9 (2023), e28. https://doi.org/10.1017/dsj.2023.27

[20] Mark Claypool. 2016. On Models for Game Input with Delay — Moving Target Selection with a Mouse. In *2016 IEEE International Symposium on Multimedia (ISM)*. IEEE, Washington, DC, USA, 575–582. https://doi.org/10.1109/ISM.2016.0125

[21] Mark Claypool. 2018. Game Input with Delay—Moving Target Selection with a Game Controller Thumbstick. *ACM Trans. Multimedia Comput. Commun. Appl.* 14, 3s, Article 57 (June 2018), 22 pages. https://doi.org/10.1145/3187288

[22] Mark Claypool, Andy Cockburn, and Carl Gutwin. 2019. Game Input with Delay: Moving Target Selection Parameters. In *Proceedings of the 10th ACM Multimedia Systems Conference* (Amherst, Massachusetts) *(MMSys '19)*. Association for Computing Machinery, New York, NY, USA, 25–35. https://doi.org/10.1145/3304109.3306232

[23] Edward R.F.W. Crossman. 1956. *The measurement of perceptual load in manual operations*. Ph. D. Dissertation. University of Birmingham.

[24] Jack Tigh Dennerlein, David B. Martin, and Christopher Hasser. 2000. Force-feedback Improves Performance for Steering and Combined Steering-targeting Tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (The Hague, The Netherlands) *(CHI '00)*. ACM, New York, NY, USA, 423–429. https://doi.org/10.1145/332040.332469

[25] Jay L. Devore. 2011. *Probability and Statistics for Engineering and the Sciences* (8th ed.). Brooks/Cole, Pacific Grove, CA. ISBN-13: 978-0-538-73352-6.

[26] Colin G. Drury. 1971. Movements with lateral constraint. *Ergonomics* 14, 2 (1971), 293–305. https://doi.org/10.1080/00140137108931246

[27] Colin G. Drury. 1994. A model for movements under intermittent illumination. *Ergonomics* 37, 7 (1994), 1245–1251. https://doi.org/10.1080/00140139408964902

[28] William R. Ferrell. 1965. Remote manipulation with transmission delay. *IEEE Transactions on Human Factors in Electronics* HFE-6, 1 (1965), 24–32. https://doi.org/10.1109/THFE.1965.6591253

[29] Leah Findlater, Joan Zhang, Jon E. Froehlich, and Karyn Moffatt. 2017. Differences in Crowdsourced vs. Lab-based Mobile and Desktop Input Performance Data. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. ACM, New York, NY, USA, 6813–6824. https://doi.org/10.1145/3025453.3025820

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

[31] Sebastian Friston, Per Karlström, and Anthony Steed. 2016. The Effects of Low Latency on Pointing and Steering Tasks. *IEEE Transactions on Visualization and Computer Graphics* 22 (2016), 1605–1615. https://doi.org/10.1109/TVCG.2015.2446467

[32] Khai-Chung Gan and Errol R. Hoffmann. 1988. Geometrical conditions for ballistic and visually controlled movements. *Ergonomics* 31, 5 (1988), 829–839. https://doi.org/10.1080/00140138808966724

[33] Brett Gilbertson. 2021. Miracast fixed with the Microsoft 4K Wireless Display Adapter! https://www.youtube.com/watch?v=bx6ujLekQV0&t=116s

[34] Jensen Harris. 2006. Giving You Fitts. https://learn.microsoft.com/en-us/archive/blogs/jensenh/giving-you-fitts

[35] Takuma Hidaka, Yusuke Sei, Naoto Nishida, Shota Yamanaka, and Buntarou Shizuki. 2023. Advanced Investigation of Steering Performance with Error-Accepting Delays. *International Journal of Human-Computer Interaction* 0, 0 (2023), 1–14. https://doi.org/10.1080/10447318.2023.2192586

[36] John W. Hill. 1976. Comparison of Seven Performance Measures in a Time-Delayed Manipulation Task. *IEEE Transactions on Systems, Man, and Cybernetics* SMC-6, 4 (1976), 286–295. https://doi.org/10.1109/TSMC.1976.5408778

[37] Errol R. Hoffmann. 1992. Fitts' Law with transmission delay. *Ergonomics* 35, 1 (1992), 37–48. https://doi.org/10.1080/00140139208967796

[38] Errol R. Hoffmann. 2009. Review of models for restricted-path movements. *International Journal of Industrial Ergonomics* 39, 4 (2009), 578–589. https://doi.org/10.1016/j.ergon.2008.02.007

[39] Errol R. Hoffmann and Shilpa Karri. 2018. Control Strategy in Movements with Transmission Delay. *Journal of Motor Behavior* 50, 4 (2018), 398–408. https://doi.org/10.1080/00222895.2017.1363700

[40] Toshiya Isomoto, Shota Yamanaka, and Buntarou Shizuki. 2022. Dwell Selection with ML-Based Intent Prediction Using Only Gaze Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 3, Article 120 (sep 2022), 21 pages. https://doi.org/10.1145/3550301

[41] Zenja Ivkovic, Ian Stavness, Carl Gutwin, and Steven Sutcliffe. 2015. Quantifying and Mitigating the Negative Effects of Local Latencies on Aiming in 3D Shooter Games. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 135–144. https://doi.org/10.1145/2702123.2702432

[42] Ricardo Jota, Albert Ng, Paul Dietz, and Daniel Wigdor. 2013. How Fast is Fast Enough?: A Study of the Effects of Latency in Direct-touch Pointing Tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. ACM, New York, NY, USA, 2291–2300. https://doi.org/10.1145/2470654.2481317

[43] Steven Komarov, Katharina Reinecke, and Krzysztof Z. Gajos. 2013. Crowdsourcing Performance Evaluations of User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. ACM, New York, NY, USA, 207–216. https://doi.org/10.1145/2470654.2470684

[44] Sergey Kulikov, I. Scott MacKenzie, and Wolfgang Stuerzlinger. 2005. Measuring the Effective Parameters of Steering Motions. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (Portland, OR, USA) *(CHI EA '05)*. ACM, New York, NY, USA, 1569–1572. https://doi.org/10.1145/1056808.1056968

[45] Sergey Kulikov and Wolfgang Stuerzlinger. 2006. Targeted Steering Motions. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI EA '06)*. ACM, New York, NY, USA, 983–988. https://doi.org/10.1145/1125451.1125640

[46] Edward Lank and Eric Saund. 2005. Sloppy Selection: Providing an Accurate Interpretation of Imprecise Selection Gestures. *Comput. Graph.* 29, 4 (Aug. 2005), 490–500. https://doi.org/10.1016/j.cag.2005.05.003

[47] Jui-Feng Lin, Colin G. Drury, Chin-Mei Chou, Yu-De Lin, and Yi-Quan Lin. 2011. Measuring Corrective Reaction Time with the Intermittent Illumination Model. In *Human-Computer Interaction. Design and Development Approaches*, Julie A. Jacko (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 397–405.

[48] Ray Lin and Colin Drury. 2010. An Application of the Intermittent Illumination Model for Measuring Individual's Corrective Reaction Time. In *The International Conference on on Applied Human Factors and Ergonomics (AHFE '10)*. AHFE, Miami, Florida. https://doi.org/10.1201/EBK1439834978-c47

[49] Ray F. Lin and Chih-Hsiang Hsu. 2014. Measuring individual corrective reaction time using the intermittent illumination model. *Ergonomics* 57, 9 (2014), 1337–1352. https://doi.org/10.1080/00140139.2014.933268

[50] Lei Liu, Jean-Bernard Martens, and Robert van Liere. 2011. Revisiting Path Steering for 3D Manipulation Tasks. *Int. J. Hum.-Comput. Stud.* 69, 3 (March 2011), 170–181. https://doi.org/10.1016/j.ijhcs.2010.11.006

[51] Shengmei Liu, Mark Claypool, Andy Cockburn, Ragnhild Eg, Carl Gutwin, and Kjetil Raaen. 2021. Datasets: Moving Target Selection with Delay. In *Proceedings of the 12th ACM Multimedia Systems Conference* (Istanbul, Turkey) *(MMSys '21)*. Association for Computing Machinery, New York, NY, USA, 320–326. https://doi.org/10.1145/3458305.3478455

[52] Michael Long and Carl Gutwin. 2018. Characterizing and Modeling the Effects of Local Latency on Game Performance and Experience. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play* (Melbourne, VIC, Australia) *(CHI PLAY '18)*. Association for Computing Machinery, New York, NY, USA, 285–297. https://doi.org/10.1145/3242671.3242678

[53] Michael Long and Carl Gutwin. 2019. Effects of Local Latency on Game Pointing Devices and Game Pointing Tasks. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300438

[54] Yuexing Luo and Daniel Vogel. 2014. Crossing-based Selection with Direct Touch Input. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. ACM, New York, NY, USA, 2627–2636. https://doi.org/10.1145/2556288.2557397

[55] I. Scott MacKenzie. 1992. Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction* 7, 1 (1992), 91–139. https://doi.org/10.1207/s15327051hci0701_3

[56] I. Scott MacKenzie and Colin Ware. 1993. Lag As a Determinant of Human Performance in Interactive Systems. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems* (Amsterdam, The Netherlands) *(CHI '93)*. ACM, New York, NY, USA, 488–493. https://doi.org/10.1145/169059.169431

[57] Iwane Maida, Tetsuro Ogi, and Tetsuya Toma. 2017. Influence of visual delay on minute pointing and push down by upper limb assuming telesurgery application. *Transactions of the Virtual Reality Society of Japan* 22, 1 (2017), 61–69. https://doi.org/10.18974/tvrsj.22.1_61

[58] M. A. Montazer, C. G. Drury, and M. H. Karwan. 1988. An optimization model for self-paced tracking on circular courses. *IEEE Transactions on Systems, Man, and Cybernetics* 18, 6 (Nov 1988), 908–916. https://doi.org/10.1109/21.23090

[59] Mahdi Nabiyouni, Siroberto Scerbo, Doug A Bowman, and Tobias Höllerer. 2017. Relative effects of real-world and virtual-world latency on an augmented reality training task: an ar simulation experiment. *Frontiers in ICT* 3 (2017), 34.

[60] Mathieu Nancel and Edward Lank. 2017. Modeling User Performance on Curved Constrained Paths. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. ACM, New York, NY, USA, 244–254. https://doi.org/10.1145/3025453.3025951

[61] Mathieu Nancel, Daniel Vogel, Bruno De Araujo, Ricardo Jota, and Géry Casiez. 2016. Next-Point Prediction Metrics for Perceived Spatial Errors. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. ACM, New York, NY, USA, 271–285. https://doi.org/10.1145/2984511.2984590

[62] Sven Nilsson. 2015. Miracast lag test. https://www.youtube.com/watch?v=szIDj4uf2hA

[63] Eunji Park and Byungjoo Lee. 2020. An Intermittent Click Planning Model. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376725

[64] Robert Pastel. 2006. Measuring the Difficulty of Steering Through Corners. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI '06)*. ACM, New York, NY, USA, 1087–1096. https://doi.org/10.1145/1124772.1124934

[65] Andriy Pavlovych and Carl Gutwin. 2012. Assessing Target Acquisition and Tracking Performance for Complex Moving Targets in the Presence of Latency and Jitter. In *Proceedings of Graphics Interface 2012* (Toronto, Ontario, Canada) *(GI '12)*. Canadian Information Processing Society, CAN, 109–116.

[66] Andriy Pavlovych and Wolfgang Stuerzlinger. 2009. The Tradeoff between Spatial Jitter and Latency in Pointing Tasks. In *Proceedings of the 1st ACM SIGCHI Symposium on Engineering Interactive Computing Systems* (Pittsburgh, PA, USA) *(EICS '09)*. Association for Computing Machinery, New York, NY, USA, 187–196. https://doi.org/10.1145/1570433.1570469

[67] Andriy Pavlovych and Wolfgang Stuerzlinger. 2009. The Tradeoff Between Spatial Jitter and Latency in Pointing Tasks. In *Proceedings of the 1st ACM SIGCHI Symposium on Engineering Interactive Computing Systems* (Pittsburgh, PA, USA) *(EICS '09)*. ACM, New York, NY, USA, 187–196. https://doi.org/10.1145/1570433.1570469

[68] Andriy Pavlovych and Wolfgang Stuerzlinger. 2011. Target Following Performance in the Presence of Latency, Jitter, and Signal Dropouts. In *Proceedings of Graphics Interface 2011* (St. John's, Newfoundland, Canada) *(GI '11)*. Canadian Human-Computer Communications Society, Waterloo, CAN, 33–40.

[69] Daniel Rakita, Bilge Mutlu, and Michael Gleicher. 2020. Effects of Onset Latency and Robot Speed Delays on Mimicry-Control Teleoperation. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction* (Cambridge, United Kingdom) *(HRI '20)*. Association for Computing Machinery, New York, NY, USA, 519–527. https://doi.org/10.1145/3319502.3374838

[70] Nicolas Rashevsky. 1959. Mathematical biophysics of automobile driving. *Bulletin of Mathematical Biophysics* 21, 4 (1959), 375–385. https://doi.org/10.1007/BF02477896

[71] Michail Schwab, Sicheng Hao, Olga Vitek, James Tompkin, Jeff Huang, and Michelle A. Borkin. 2019. Evaluating Pan and Zoom Timelines and Sliders. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300786

[72] Ransalu Senanayake and Ravindra S. Goonetilleke. 2016. Pointing Device Performance in Steering Tasks. *Perceptual and Motor Skills* 122, 3 (2016), 886–910. https://doi.org/10.1177/0031512516649717

[73] Ransalu Senanayake, Errol R. Hoffmann, and Ravindra S. Goonetilleke. 2013. A model for combined targeting and tracking tasks in computer applications. *Experimental Brain Research* 231, 3 (01 Nov 2013), 367–379. https://doi.org/10.1007/s00221-013-3700-4

[74] T.B. Sheridan and William R. Ferrell. 1963. Remote Manipulative Control with Transmission Delay. *IEEE Transactions on Human Factors in Electronics* HFE-4, 1 (1963), 25–29. https://doi.org/10.1109/THFE.1963.231283

[75] Richard H.Y. So, Germen K.M. Chung, and Ravindra S. Goonetilleke. 1999. Target-Directed Head Movements in a Head-Coupled Virtual Environment: Predicting the Effects of Lags Using Fitts' Law. *Human Factors* 41, 3 (1999), 474–486. https://doi.org/10.1518/001872099779611067

[76] 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. https://doi.org/10.1016/j.ijhcs.2004.09.001

[77] Gregory P. Starr. 1979. A Comparison of Control Modes for Time-Delayed Remote Manipulation. *IEEE Transactions on Systems, Man, and Cybernetics* 9, 4 (1979), 241–246. https://doi.org/10.1109/TSMC.1979.4310186

[78] Ahmed N. Sulaiman and Patrick Olivier. 2008. Attribute Gates. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology* (Monterey, CA, USA) *(UIST '08)*. ACM, New York, NY, USA, 57–66. https://doi.org/10.1145/1449715.1449726

[79] Minghui Sun, Xiangshi Ren, and Xiang Cao. 2010. Effects of Multimodal Error Feedback on Human Performance in Steering Tasks. *Journal of Information Processing* 18 (2010), 284–292. https://doi.org/10.2197/ipsjjip.18.284

[80] Erum Tanvir, Andrea Bunt, Andy Cockburn, and Pourang Irani. 2011. Improving cascading menu selections with adaptive activation areas. *International Journal of Human-Computer Studies* 69, 11 (2011), 769 – 785. https://doi.org/10.1016/j.ijhcs.2011.06.005

[81] Erum Tanvir, Jonathan Cullen, Pourang Irani, and Andy Cockburn. 2008. AAMU: Adaptive Activation Area Menus for Improving Selection in Cascading Pull-down Menus. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. ACM, New York, NY, USA, 1381–1384. https://doi.org/10.1145/1357054.1357270

[82] Robert J. Teather, Andriy Pavlovych, Wolfgang Stuerzlinger, and I. Scott MacKenzie. 2009. Effects of Tracking Technology, Latency, and Spatial Jitter on Object Movement. In *Proceedings of the 2009 IEEE Symposium on 3D User Interfaces (3DUI '09)*. IEEE Computer Society, Washington, DC, USA, 43–50. https://doi.org/10.1109/3DUI.2009.4811204

[83] Namal Thibbotuwawa, Ravindra S. Goonetilleke, and Errol R. Hoffmann. 2012. Constrained Path Tracking at Varying Angles in a Mouse Tracking Task. *Human Factors* 54, 1 (2012), 138–150. https://doi.org/10.1177/0018720811424743

[84] Namal Thibbotuwawa, Errol R. Hoffmann, and Ravindra S. Goonetilleke. 2012. Open-loop and feedback-controlled mouse cursor movements in linear paths. *Ergonomics* 55, 4 (2012), 476–488. https://doi.org/10.1080/00140139.2011.644587

[85] Himari Tochioka, Haruka Ikeda, Tomohiko Hayakawa, and Masatoshi Ishikawa. 2019. Effects of Latency in Visual Feedback on Human Performance of Path-Steering Tasks. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 65, 2 pages. https://doi.org/10.1145/3359996.3364726

[86] Huawei Tu, Susu Huang, Jiabin Yuan, Xiangshi Ren, and Feng Tian. 2019. Crossing-Based Selection with Virtual Reality Head-Mounted Displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. ACM, New York, NY, USA, Article 618, 14 pages. https://doi.org/10.1145/3290605.3300848

[87] Hiroki Usuba, Shota Yamanaka, and Homei Miyashita. 2019. Comparing Lassoing Criteria and Modeling Straight-Line and One-Loop Lassoing Motions Considering Criteria. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces* (Daejeon, Republic of Korea) *(ISS '19)*. Association for Computing Machinery, New York, NY, USA, 181–191. https://doi.org/10.1145/3343055.3359707

[88] Ruben Geert van den Berg. 2022. SPSS TUTORIALS – Effect Size - A Quick Guide. https://www.spss-tutorials.com/effect-size/

[89] Matthew Warburton, Mark Mon-Williams, Faisal Mushtaq, and J Morehead. 2022. Measuring motion-to-photon latency for sensorimotor experiments with virtual reality systems. *Behavior research methods* 55, 7 (10 2022), 3658–3678. https://doi.org/10.3758/s13428-022-01983-5

[90] Colin Ware and Ravin Balakrishnan. 1994. Reaching for Objects in VR Displays: Lag and Frame Rate. *ACM Trans. Comput.-Hum. Interact.* 1, 4 (dec 1994), 331–356. https://doi.org/10.1145/198425.198426

[91] Jannik Wiese and Niels Henze. 2023. Predicting Mouse Positions Beyond a System's Latency Can Increase Throughput and User Experience in Linear Steering Tasks. In *Proceedings of Mensch Und Computer 2023* (Rapperswil, Switzerland) *(MuC '23)*. Association for Computing Machinery, New York, NY, USA, 101–115. https://doi.org/10.1145/3603555.3603556

[92] Jacob O. Wobbrock and Krzysztof Z. Gajos. 2008. Goal Crossing with Mice and Trackballs for People with Motor Impairments: Performance, Submovements, and Design Directions. *ACM Trans. Access. Comput.* 1, 1, Article 4 (May 2008), 37 pages. https://doi.org/10.1145/1361203.1361207

[93] Haijun Xia, Ricardo Jota, Benjamin McCanny, Zhe Yu, Clifton Forlines, Karan Singh, and Daniel Wigdor. 2014. Zero-Latency Tapping: Using Hover Information to Predict Touch Locations and Eliminate Touchdown Latency. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST '14)*. Association for Computing Machinery, New York, NY, USA, 205–214. https://doi.org/10.1145/2642918.2647348

[94] Yahoo Japan Corporation. 2023. Yahoo! Crowdsourcing. https://crowdsourcing.yahoo.co.jp

[95] Shota Yamanaka. 2019. Steering Performance with Error-accepting Delays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. ACM, New York, NY, USA, Article 570, 9 pages. https://doi.org/10.1145/3290605.3300800

[96] Shota Yamanaka. 2021. Comparing Performance Models for Bivariate Pointing through a Crowdsourced Experiment. In *Human-Computer Interaction – INTERACT 2021*. Springer International Publishing, Gewerbestr, Switzerland, 76–92. https://doi.org/10.1007/978-3-030-85616-8_6

[97] Shota Yamanaka. 2021. Utility of Crowdsourced User Experiments for Measuring the Central Tendency of User Performance to Evaluate Error-Rate Models on GUIs. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* 9, 1 (Oct. 2021), 155–165.

[98] Shota Yamanaka and Homei Miyashita. 2016. Modeling the Steering Time Difference Between Narrowing and Widening Tunnels. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. ACM, New York, NY, USA, 1846–1856. https://doi.org/10.1145/2858036.2858037

[99] Shota Yamanaka and Homei Miyashita. 2016. Scale Effects in the Steering Time Difference Between Narrowing and Widening Linear Tunnels. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction* (Gothenburg, Sweden) *(NordiCHI '16)*. ACM, New York, NY, USA, Article 12, 10 pages. https://doi.org/10.1145/2971485.2971486

[100] Shota Yamanaka and Homei Miyashita. 2019. Modeling Pen Steering Performance in a Single Constant-Width Curved Path. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces* (Daejeon, Republic of Korea) *(ISS '19)*. Association for Computing Machinery, New York, NY, USA, 65–76. https://doi.org/10.1145/3343055.3359697

[101] Shota Yamanaka and Wolfgang Stuerzlinger. 2019. Modeling Fully and Partially Constrained Lasso Movements in a Grid of Icons. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. ACM, New York, NY, USA, Article 120, 12 pages. https://doi.org/10.1145/3290605.3300350

[102] Shota Yamanaka and Wolfgang Stuerzlinger. 2020. Necessary and Unnecessary Distractor Avoidance Movements Affect User Behaviors in Crossing Operations. *ACM Trans. Comput.-Hum. Interact.* 27, 6, Article 44 (nov 2020), 31 pages. https://doi.org/10.1145/3418413

[103] Shota Yamanaka, Wolfgang Stuerzlinger, and Homei Miyashita. 2017. Steering Through Sequential Linear Path Segments. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. ACM, New York, NY, USA, 232–243. https://doi.org/10.1145/3025453.3025836

[104] Shota Yamanaka, Wolfgang Stuerzlinger, and Homei Miyashita. 2018. Steering Through Successive Objects. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 603, 13 pages. https://doi.org/10.1145/3173574.3174177

[105] Shota Yamanaka and Hiroki Usuba. 2020. Rethinking the Dual Gaussian Distribution Model for Predicting Touch Accuracy in On-Screen-Start Pointing Tasks. *Proc. ACM Hum.-Comput. Interact.* 4, ISS, Article 205 (Nov. 2020), 20 pages. https://doi.org/10.1145/3427333

[106] Shota Yamanaka, Hiroki Usuba, Wolfgang Stuerzlinger, and Homei Miyashita. 2022. The Effectiveness of Path-Segmentation for Modeling Lasso Times in Width-Varying Paths. *Proc. ACM Hum.-Comput. Interact.* 6, ISS, Article 584 (nov 2022), 20 pages. https://doi.org/10.1145/3567737

[107] Shota Yamanaka, Hiroki Usuba, Haruki Takahashi, and Homei Miyashita. 2023. Predicting Success Rates in Steering Through Linear and Circular Paths by the Servo-Gaussian Model. *International Journal of Human-Computer Interaction* 40 (2023), 1–14. https://doi.org/10.1080/10447318.2023.2192586

[108] Eun Kyoung Yang and Jee Hyun Lee. 2020. Cognitive impact of virtual reality sketching on designers' concept generation. *Digital Creativity* 31, 2 (2020), 82–97. https://doi.org/10.1080/14626268.2020.1726964

[109] Xing-Dong Yang, Pourang Irani, Pierre Boulanger, and Walter F. Bischof. 2009. A Model for Steering with Haptic-Force Guidance. In *Proceedings of the 12th IFIP International Conference on Human-Computer Interaction (INTERACT '09)*. Springer Berlin Heidelberg, Berlin, Heidelberg, 465–478. https://doi.org/10.1007/978-3-642-03658-3_51

[110] Difeng Yu, Hai-Ning Liang, Xueshi Lu, Kaixuan Fan, and Barrett Ens. 2019. Modeling Endpoint Distribution of Pointing Selection Tasks in Virtual Reality Environments. *ACM Trans. Graph.* 38, 6, Article 218 (Nov. 2019), 13 pages. https://doi.org/10.1145/3355089.3356544

[111] Xiaolei Zhou and Xiangshi Ren. 2010. An investigation of subjective operational biases in steering tasks evaluation. *Behaviour & Information Technology* 29, 2 (2010), 125–135. https://doi.org/10.1080/01449290701773701

[112] Xiaolei Zhou, Xiangshi Ren, and Hue YUi. 2008. Effect of Start Position on Human Performance in Steering Tasks. In *2008 International Conference on Computer Science and Software Engineering*, Vol. 2. IEEE, Washington, DC, USA, 1098–1101. https://doi.org/10.1109/CSSE.2008.1310