# The Effects of Task Factors on the Multi-Directional Tapping Task

Yuhwa Hong [a,1] (yuhwahong@kaist.ac.kr; 0009-0002-6622-304X),
Haejun Kim [a,1] (vedonefor@kaist.ac.kr; 0009-0001-0717-6776),
Jihae Yu [b] (vievanitas@kangwon.ac.kr; 0009-0002-4366-1776),
Heedo Shin [b] (gojob123@naver.com; 0009-0003-3229-8390),
Xiaoqun Yu [c] (xiaoqunyu@seu.edu.cn; 0000-0001-8795-5237),
Shuping Xiong [a,*] (shupingx@kaist.ac.kr; 0000-0003-1549-515X),
Woojoo Kim [b,*] (woojoo.kim@kangwon.ac.kr; 0000-0001-6203-7309)

[a] Department of Industrial and Systems Engineering, Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea;
[b] Kangwon National University, 1 Kangwondaehak-gil, Chuncheon-si, Gangwon-do 24341, Republic of Korea;
[c] Department of Mechanical and Industrial Design, Southeast University, 2 Southeast University Road, Jiangning District, Nanjing 211189, China

**ARTICLE HISTORY**
Compiled September 29, 2025

**ABSTRACT**
In Fitts' law research for pointing and selecting tasks, the multi-directional tapping task from ISO/TS 9241-411 has been widely used as a standard task. However, the ISO standard does not describe several task factors in sufficient detail, leading researchers to interpret and apply them independently in various ways. This study aims to investigate the effects of four task factors for multi-directional tapping task: error highlight, hover highlight, target shape, and number of targets. The experimental results indicated that all investigated task factors had significant effects on user performance and behavior, as reflected in both eye movement and mouse movement patterns. Our research findings can be helpful for researchers who want to evaluate the efficiency and effectiveness of interaction techniques and input devices for graphical user interfaces as a multi-directional tapping task protocol.

**KEYWORDS**
Multi-directional tapping task; Target selection; Task factors; Fitts' law

## 1. Introduction

In the field of Human-Computer Interaction (HCI), the evaluation of efficiency and effectiveness of interaction techniques and input devices for graphical user interfaces (GUIs) has been a crucial area of research (Cassidy, Read, & MacKenzie, 2019; Cockburn, Ahlström, & Gutwin, 2012). One of the most widely used methods for this evaluation is the multi-directional tapping task, which is standardized in ISO/TS 9241-411 (ISO, 2012) (previously known as ISO 9241-9 (ISO, 2000)). This task is also known

*Corresponding author
(Shuping Xiong) Email: shupingx@kaist.ac.kr; (Woojoo Kim) Email: woojoo.kim@kangwon.ac.kr
[1] The authors share the first authorship.

as 2D Fitts' law task, and has been extensively utilized in numerous studies to assess various performance aspects of user interaction with GUIs (Cheong, Kim, Park, & Park, 2011; Pino, Tzemis, Ioannou, & Kouroupetroglou, 2013; Roig-Maimó, MacKenzie, Manresa-Yee, & Varona, 2018; Wobbrock, Jansen, & Shinohara, 2011) and even with 3D UIs in the virtual environment (Batmaz, Yu, Liang, & Stuerzlinger, 2022; Li et al., 2024; Yu, Liang, Lu, Fan, & Ens, 2019).

Fitts' law quantifies how movement time is influenced by target distance and size during rapid pointing movements (Fitts, 1954). Fitts' original paradigm utilized two rectangular targets to assess one-dimensional horizontal movement, but this classic approach has become outdated due to its limitation in accounting for how movement direction affects pointing performance. To address this shortcoming, the ISO Technical Specification (ISO/TS) advocates for a circular target configuration, known as the multi-directional tapping task (ISO, 2012; Soukoreff & MacKenzie, 2004). In this task, targets are arranged in a circular pattern on a display or surface. Participants must tap or click these targets in a specific sequence, typically alternating between opposite sides of the circle. The multi-directional tapping task provides a structured approach to measuring user performance in target acquisition tasks, which are fundamental to many computer interactions.

However, despite its widespread use and inclusion in ISO/TS, the description of the multi-directional tapping task in the ISO/TS 9241-411 document is notably brief and lacks detailed specifications for various task factors. This ambiguity has led to a situation where researchers have been compelled to make independent decisions regarding the implementation of these factors in their experiments. Consequently, there is significant variation across studies (Cassidy et al., 2019; Cockburn & Brock, 2006; Matsuyama & Karashima, 2017) in how certain aspects of the task are presented and executed. Some of the key task factors that have been inconsistently applied across different studies include (More details are described in '2. Related Work'):

- **Error Highlight:** The method used to inform participants when they have made an error in target selection. Indication of when the selection occurs outside of the target varies across the studies. Examples include using a beep sound (Cuaresma & Mackenzie, 2017), changing the target color (S. Kim, Lee, Gemert, & Oulasvirta, 2020), or combining visual and auditory cues like a flashing red target with a ding sound (Wobbrock, Jansen, & Shinohara, 2011). Conversely, no highlighting is provided for errors in some cases (H. J. Chen, Lin, & Lin, 2019).
- **Hover Highlight:** The visual feedback provided to show when the cursor is within the target area. Some studies have employed feedback mechanisms that indicate the target's state through color changes when the cursor touches the target (Herring, Trejo, & Hallbeck, 2010; Roig-Maimó et al., 2018), while others did not.
- **Target Shape:** The geometric form of the targets used in the task. Despite the ISO/TS mentioning the use of square targets as an example, very few studies (Matsuyama & Karashima, 2017; Park, Hong, & Lee, 2012) utilize square targets, while the vast majority of studies use circular targets (Avsar, Fischer, & Rodden, 2016; W. Kim & Xiong, 2022; Mackenzie & Teather, 2012; Pandey & Arif, 2022; Rajanna & Hammond, 2018).
- **Number of Targets:** The quantity of targets presented in the task layout. Various number of targets has been applied to the studies, from less than 10 to more than 20 (Boritz, Booth, human performance, & undefined 1991, 1991; Gori

& Bellut, 2023; Pandey & Arif, 2022; Wang, Yu, Qin, Li, & Shi, 2013).

These variations in task implementation raise important questions about the comparability and generalizability of results across different studies. Surprisingly, despite the potential impact of these factors on user performance, there has been limited research investigating how these diverse task factors might significantly influence the outcomes of the multi-directional tapping task.

Given this gap in our understanding, there is a clear need for a more standardized protocol for the multi-directional tapping task. Such standardization would not only enhance the reliability and comparability of results across different studies but also provide a more solid foundation for future research in this area. To address this need, our study aims to investigate the effects of aforementioned four specific task factors on user performance and behavior in the multi-directional tapping task.

The main contributions of our study are as follows:

- This study provides a comprehensive investigation on the effects of four task factors in multi-directional tapping tasks, revealing their significant effects on user performance and behavior.
- We found that the four task factors significantly influenced user performance and behavior in the multi-directional tapping task. When error highlight was provided, the error rate decreased but movement time increased. When hover highlight was given, the throughput decreased. Square targets induced lower movement time compared to circular targets, and a lower number of targets increased throughput. Findings underscore the importance for researchers to carefully consider these task factors when designing future studies involving this task.
- This research has the potential to improve the consistency and validity of future studies in HCI, particularly those focusing on evaluation of input devices and interaction techniques. We present specific recommendations for implementing the properties of multi-directional tapping tasks to assist researchers in designing more standardized and effective experimental protocols for future studies.

## 2. Related Work

### *2.1. Fitts' Law and Effective Parameters*

Fitts' law (Fitts, 1954) is a predictive model of human movement which states that the time to acquire a target is a function of the distance to and size of the target (Equation 1), where $MT$ is the movement time, $a$ and $b$ are empirically determined constants, $A$ is the target amplitude, and $W$ is the target width. The Shannon formulation of the index of difficulty ($ID$) (MacKenzie, 1989; Soukoreff & MacKenzie, 2004) is most widely used in HCI. Fitts' law has been widely applied in the field of HCI, particularly in the design and evaluation of UIs, input devices, and interaction techniques (MacKenzie & Buxton, 1992).

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

While the original formulation of Fitts' law uses nominal target amplitude ($A$) and target width ($W$), researchers have introduced the concept of effective parameters to account for the observed variability in human performance and to improve the law's predictive power (Soukoreff & MacKenzie, 2004). Effective target width ($W_e$) is a measure that takes into account the variability in endpoint positions when users attempt to hit a target. It is typically calculated as: $W_e = 4.133 \times SD_x$ , where $SD_x$ is the standard deviation of the selection coordinates measured along the task axis. Effective target amplitude ($A_e$) is the mean distance traveled by the pointer. By incorporating these effective parameters, the effective index of difficulty ($ID_e$) can be calculated as (Equation 2):

$$ID_e = \log_2 \left( \frac{A_e}{W_e} + 1 \right) \tag{2}$$

Throughput ($TP$) is a comprehensive measure that captures user performance in Fitts' law experiments, considering both speed and accuracy. One of its key advantages is that it remains constant across various speed-accuracy trade-offs (MacKenzie & Isokoski, 2008), facilitating easy comparisons of performance under different conditions. The (effective) throughput can be calculated as (Equation 3):

$$TP = \frac{ID_e}{MT} \tag{3}$$

### *2.2. The Multi-directional Tapping Task*

From choosing an icon on a desktop to navigating complex web applications, selections are the building blocks of our digital experiences. In computer-based interactions, selecting—through pointing, clicking, or tapping—stands as a fundamental pillar of how users interact with digital interfaces to achieve their desired outcomes. In this study, selecting is considered as a target acquisition process that can be carried out through different interaction methods, including pointing, clicking, and tapping. While these interactions differ in their physical execution—pointing involves moving a pointer or contact point toward the target, clicking refers to pressing and releasing a button, and tapping refers to briefly touching a touch-sensitive surface—they all serve as mechanisms for confirming a target choice within the broader selecting action.

Researchers have used various methods to study selection tasks. Some have observed natural web browsing behavior (Sundar, Bellur, Oh, Xu, & Jia, 2014; Zhai, Smith, & Selker, 1997), while others have designed experiments with users targeting random objects on a screen (Roig-Maimó, MacKenzie, Manresa-Yee, & Varona, 2017). Although these approaches have provided valuable insights, a need for standardization emerged as HCI research evolved. Researchers required a consistent, replicable method to study selection tasks across different contexts and input devices, which led to the widespread adoption of the multi-directional tapping task, as specified in the ISO/TS 9241-411 (ISO, 2012). Table 1 shows the full description of the multi-directional tapping task from ISO/TS 9241-411 (ISO, 2012).

**Table 1.** The full description of the multi-directional tapping task from ISO/TS 9241-411 (ISO, 2012)

| **Annex B. Testing of efficiency and effectiveness – B.6.2.2 Multi-directional tapping test** |
|---|
| This test can be used to evaluate: pointing movements in many different directions. Its applications include a) repositioning a cursor at different areas on the screen, b) cell selection in a spreadsheet, and c) selecting randomly located icons.<br>**Test procedure:** The subject is required to move the cursor across a circle to sequentially numbered targets (see Figure 1). The targets (for example: squares) should be equally spaced around the circumference of the circle. The targets should be arranged so that the movements are nearly equal to the diameter of the circle. The targets to which the subject should advance should be highlighted. Each test session starts after the subject points to the topmost target and ends when the sequence is completed (at the topmost target).<br>This test should be conducted with a range of difficulties. That is, the size of the circle and thus the distance between the target squares should be varied between trials, as long as all subjects have the same test conditions.<br>The results should be calculated according to Equations 2 and 3. |

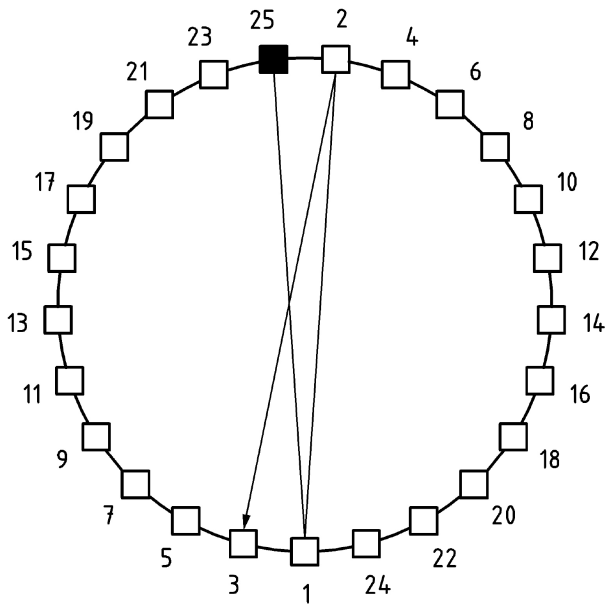



**Figure 1.** A figure describing the multi-directional tapping task from ISO/TS 9241-411 (ISO, 2012).

The multi-directional tapping task has become a cornerstone in HCI research, finding applications across a wide spectrum of studies. Researchers have employed this versatile task for various purposes, such as comparing the effectiveness of different input devices (Beelders & Blignaut, 2012; Cheong et al., 2011; Matsuyama & Karashima, 2017; Natapov & MacKenzie, 2010; Rajanna & Hammond, 2018), validating new experimental protocols (Gori & Bellut, 2023), and investigating how display sizes affect user performance (Browning & Teather, 2014; Wang et al., 2013). Beyond these applications, studies have also used it to examine unconventional input modes such as gaze, foot-based interfaces, and gesture controls (Dube, Ren, Limerick, Mackenzie, & Arif, 2022; José & Lopes, 2015; Katsuragawa, Pietroszek, Wallace, & Lank, 2016; Velloso, Alexander, Bulling, & Gellersen, 2015; Zhang & MacKenzie, 2007). Interestingly, despite being originally designed for evaluation of 2D interactions, the multi-directional tapping task has found its way into virtual reality (VR) research (Amini, Stuerzlinger, Teather, & Batmaz, 2025; Batmaz et al., 2022; W. Kim & Xiong, 2022; Li et al., 2024; Yu et al., 2019). This adoption stems largely from the lack of a standardized 3D Fitts' law task, making it a go-to choice for researchers in this emerging field.

### *2.3. Factors in the Multi-directional Tapping Task*

Despite widespread use of the multi-directional tapping task, the brevity of the task's description in the standard has led to significant variability in experimental designs across studies (Gillan, Holden, Adam, Rudisill, & Magee, 1990; Pino et al., 2013; Roig-Maimó et al., 2018). While Fitts' law studies typically provide clear explanations for key parameters like $W$, $A$, and $ID$—parameters known to directly influence performance, other crucial experimental factors often lack such clarity. Some additional task elements, including target shape, number of targets, and feedback during task execution, may significantly impact user performance and experimental outcomes.

Research on the effects of these factors is limited. Some studies have explored the significance of target height or aspect ratio in Fitts' law tasks (Hoffmann & Sheikh, 1994; Ko et al., 2020; Zhang, Zha, & Feng, 2012), leading to the recognition of square targets as meaningful shapes in user interfaces (Beelders & Blignaut, 2012; Boritz et al., 1991; Matsuyama & Karashima, 2017; Park et al., 2012). However, there's a lack of comprehensive research comparing rectangular shapes to other geometric forms. Sheikh and Hoffmann (Sheikh & Hoffmann, 1994) found significant differences in movement time among squares, circles, diamonds, and triangles, although this old study employed Fitts' original one-dimensional paradigm. Yet circular targets are used dominantly in studies (Avsar et al., 2016; Bernardos, Gómez, & Casar, 2016; Kouroupetroglou et al., 2012; McArthur, Castellucci, & MacKenzie, 2009; Natapov & MacKenzie, 2010), often without clear justification beyond citing the ISO/TS.

The number of targets used in experiments varies widely, ranging from fewer than 10 to over 20. Table 2 shows how various target numbers were used in previous studies. However, first 3 trials (first to third target selections) were considered practice and therefore excluded from the analysis in some research (Boritz et al., 1991; Hertzum & Hornbæk, 2013; Natapov & MacKenzie, 2010; Wobbrock, Jansen, & Shinohara, 2011). While Fitts' law primarily defines the $ID$ using $W$ and $A$, it is important to note that increasing the number of targets (effectively increasing the number of tasks) is likely to affect overall task difficulty by increasing the workload. Despite this potential impact, most studies fail to provide explanations for their chosen number of targets. For instance, Weiss et al. (Weiss, Tang, Williams, & Stirling, 2024) opted for 16 targets in their study, a notable reduction from the 25 targets used in previous research, but they did not provide a clear rationale for this choice.

**Table 2.** Summary of various target numbers used in the previous research

| Number of Targets | References |
|---|---|
| 9 targets | Ortega and Nigay (2009), Cuaresma and Mackenzie (2017), Pandey and Arif (2022) |
| 11 targets | Dube et al. (2022) |
| 12 targets | Mackenzie and Teather (2012), Yuan, Calic, Fernando, and Kondoz (2013), Katsuragawa et al. (2016), Bernardos et al. (2016), H. J. Chen et al. (2019) |
| 13 targets | Velloso et al. (2015), Avsar et al. (2016), Ramcharitar and Teather (2017), Rajanna and Hammond (2018), Li et al. (2024) |
| 14 targets | Browning and Teather (2014), José and Lopes (2015) |
| 15 targets | McArthur et al. (2009), Herring et al. (2010), Pino et al. (2013), Draghici, Batkin, Bolic, and Chapman (2014), Matsuyama and Karashima (2017), Cassidy et al. (2019), S. Kim et al. (2020), Li, Sarcar, Kim, Tu, and Ren (2022), Roig-Maimó and Mas-Sansó (2023) |
| 16 targets | Zhang and MacKenzie (2007), Cheong et al. (2011), Beelders and Blignaut (2012), Kouroupetroglou et al. (2012), Janthanasub and Meesad (2015) |
| 17 targets | Cockburn and Brock (2006) |
| 20 targets | Roig-Maimó et al. (2018) |
| 21 targets | Yu et al. (2019), Gori and Bellut (2023) |
| 25 targets | Park et al. (2012) |
| 26 targets | Wang et al. (2013) |

The issue of inconsistency extends to the feedback provided during task performance. While numerous studies have shown that feedback significantly influences task performance and user behavior (Fogg & Nass, 1997; Te'eni, 1992), there is a lack of research specifically examining the effects of feedback in the multi-directional tapping task. Interestingly, Batmaz and Stuerzlinger (Batmaz & Stuerzlinger, 2021) reported that in VR multi-directional tapping tasks, performance decreased in terms of time and throughput when high-frequency pitch error sound feedback was provided, suggesting that the relationship between feedback and performance may be complex. Despite this knowledge gap, researchers have implemented various forms of feedback in their studies. Some have used visual cues, such as highlighting the target when the cursor enters it (Herring et al., 2010; Roig-Maimó et al., 2018), while others have displayed the next target to be selected after the current one (Bateman, Mandryk, Gutwin, & Xiao, 2013; W. Kim & Xiong, 2022). When participants miss a target, researchers have generally used one of or both visual and sound effects to indicate error (Cuaresma & Mackenzie, 2017; S. Kim et al., 2020; MacKenzie, Kauppinen, & Silfverberg, 2001; Wobbrock, Jansen, & Shinohara, 2011), whereas others offer no indication (H. J. Chen et al., 2019; Hertzum & Hornbæk, 2013).

The inconsistent use of various task conditions in such studies likely stems from the lack of specific guidelines in the description from ISO/TS 9241-411. Given that even slight differences in experimental conditions can lead to variations in results, there's a clear need for more detailed investigation into these factors. Therefore, this study aims to explore the effect of factors in the multi-directional tapping task that previous research has not adequately addressed, by analyzing performance and behavioral outcomes.

## 3. Experiment

### 3.1. Participants

We recruited 38 adults (24 females and 14 males; age: M = 22.32, SD = 4.34) with normal or corrected-to-normal vision, all of whom were right-handed. The majority of participants used a PC either 3 to 6 times per week (16 participants) or daily (16 participants). 31 participants reported themselves as intermediate regarding their proficiency in using a PC. Additionally, 21 participants reported using a mouse daily, and 22 participants considered themselves proficient in its use. All participants gave informed consent for the experiment protocol, which was approved by the Institutional Review Board at Kangwon National University (IRB NO.: KWNUIRB-2024-09-006-001).

### 3.2. Apparatus and Experimental Setup

Figure 2 shows the experimental setup. The experiment was conducted using a PC equipped with an Intel Core i7-14700 processor, 64GB of RAM, and an NVIDIA GeForce RTX 4070 Ti SUPER graphics card, running Windows 11 Pro. An LG UltraGear 24GS60F monitor (24-inch, 1920 × 1080 resolution, 180Hz refresh rate) was used for display, and a Tobii Pro Spark eye tracker (60Hz sampling rate) was attached to the bottom of the monitor to record eye movement data. Participants used a Logitech M90 optical wired mouse (1000 DPI) to control the cursor. The mouse without a DPI switch was deliberately chosen to prevent any accidental changes in sensitivity during the experiment. Participants were seated comfortably at a desk, facing the monitor and holding the mouse in their right hand to perform the given tasks. No keyboard was present on the desk. The experimental program was developed using Unity 2022.3.42f1.

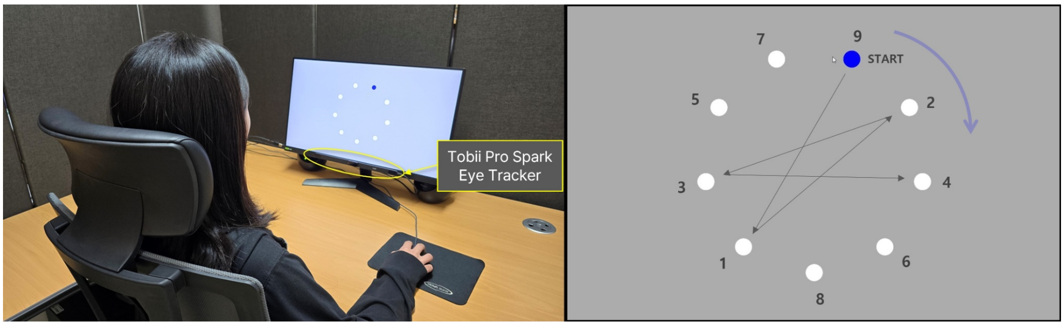
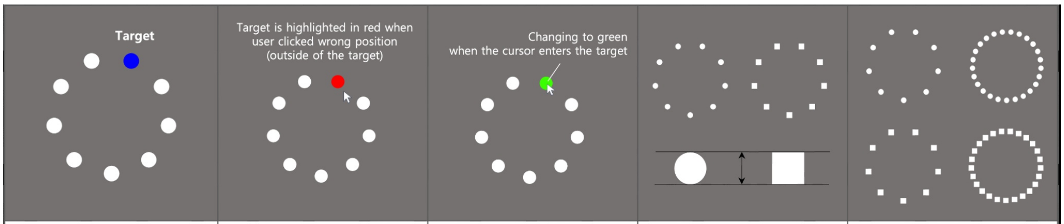



**Figure 2.** The experimental setup (left) and task screen illustrating the detailed sequence and direction of tapping (right).

### 3.3. Experimental Design

This study employed a 2 × 2 × 2 × 2 within-subjects design with the independent variables and levels shown in Figure 3, which will hereafter be collectively referred to as the task factors for clarity.




| Task Factors | 1) Error Highlight (**EH**) | 2) Hover Highlight (**HH**) | 3) Target Shape (**TS**) | 4) Number of Targets (**TN**) |
|---|---|---|---|---|
| **Levels** | • With ($EH_{with}$)<br>• Without ($EH_{without}$) | • With ($HH_{with}$)<br>• Without ($HH_{without}$) | • Circular ($TS_{circular}$)<br>• Square ($TS_{square}$) | • 9 ($TN_9$)<br>• 25 ($TN_{25}$) |

**Figure 3.** Four factors of the multi-directional tapping task investigated in this study.

- **Error Highlight (*EH*)**: 2 conditions were tested: *with* and *without* error highlight. When participants selected outside of targeting object, the error highlight was given by changing the missed target to red for a 0.1-second and playing a short beep sound, similar to the setup used in *FittsStudy* (Wobbrock, Shinohara, & Jansen, 2011).
- **Hover Highlight (*HH*)**: 2 conditions were tested: *with* and *without* hover highlight. When the cursor was located inside the target, the hover highlight changed the target color to green.
- **Target Shape (*TS*)**: 2 conditions were tested: targets with *circular* and *square* shapes. The $W$ determined the diameter of circular targets and the side length of square targets.
- **Number of Targets (*TN*)**: 2 conditions were tested: *9* and *25* targets. Notably, we selected numbers close to the minimum and maximum used in the literature.

To investigate the effect of task factors across a wide range of $ID$ values in a Fitts' law test, we included combinations of 3 target widths (20, 50, and 110 pixels) and 3 target amplitudes (320, 640, and 960 pixels), resulting in 9 distinct $ID$ conditions (1.97, 2.77, 2.89, 3.28, 3.79, 4.09, 4.32, 5.04, and 5.61). These 9 ID levels were adapted from previous studies involving multi-directional tapping tasks, reflecting the commonly used range of approximately 1 to 6 (Oehl, Sutter, & Ziefle, 2007; Velloso et al., 2015; Wobbrock, Jansen, & Shinohara, 2011; Wurth & Hargrove, 2014).

This design led to a total of 2 error highlight status (*with* and *without*) $\times$ 2 hover highlight status (*with* and *without*) $\times$ 2 target shapes (*circular* and *square*) $\times$ (*9* and *25*) targets $\times$ 3 target widths $\times$ 3 target amplitudes = 2448 data points per participant. We included all trial data in the analysis, even those containing errors, to ensure a comprehensive evaluation with the calculation of effective width ($W_e$).

### *3.4. Experimental Procedure*

After receiving a detailed explanation of the experiment and providing informed consent, participants engaged in a practice session. This session consisted of a series of multi-directional tapping tasks with a $W = 50$ pixels and $A = 640$ pixels, involving one round of trials for each of the 16 task conditions (combinations of 4 task factors), presented in a randomized order. Subsequently, participants proceeded to the main experimental session, where they performed the multi-directional tapping task once for each of the $2 \times 2 \times 2 \times 2 \times 3 \times 3 = 144$ test conditions with instructions to perform pointing movements. Additionally, they were instructed to perform the tasks as fast and precise as possible under the accuracy-emphasis condition of the manipulating strategies (Batmaz & Stuerzlinger, 2022; Olafsdottir, Guiard, Rioul, & Perrault, 2012). The presentation order was structured as follows: first, the 16 task conditions were randomized. Within each task condition, the 9 ID conditions were then presented in a random order. This approach, as opposed to fully randomizing all 144 test conditions, was adopted to facilitate the collection of data on participants' behavioral changes when performing tasks consecutively within the same task condition.

Upon completion of all test conditions, participants engaged in a short interview. During this interview, they were asked to answer whether they perceived each of the 4 task conditions as influential on their task performance, and to provide rationales for their perceptions. For example, participants were asked, *"Did you feel a difference depending on whether error highlighting was present or not?"* and *"If you felt a difference, what do you think caused it?"* These questions were posed in a semi-structured format, allowing participants to elaborate freely on their responses. The experiment took around 40-50 minutes to complete, and each participant received a monetary reward of KRW 30,000 ($\sim$ USD 22) for completing the experiment.

### *3.5. Data Analysis*

For performance measures: movement time ($MT$), error rate ($ER$), and effective throughput ($TP$) (Mackenzie, 2015) were collected and arranged for the analysis. All trials including those with errors were considered for analysis like in earlier studies on pointing and crossing (Kasahara, Oba, Yamanaka, Stuerzlinger, & Miyashita, 2023; MacKenzie & Isokoski, 2008). The inclusion of all trials allows for the calculation of not only $ER$ but also $W_e$, which provides a measure of the overall effective task difficulty. To identify outliers, the interquartile range (IQR) method was employed, defining potential outliers as participants whose ER exceeded 1.5 times the IQR. However, no such outliers were detected, so the analysis proceeded with the complete dataset without any participant exclusions.

Furthermore, to gain a deeper understanding of the underlying reasons for changes in user performance and the effect of task factors on user behavior, we analyzed cursor and eye movement measures. For cursor movement, we examined accuracy measures for evaluating computer pointing devices (MacKenzie et al., 2001), including target re-entry, task axis crossing, movement direction change, orthogonal direction change, movement variability, movement error, and movement offset. These measures provided insights into cursor movement patterns. Eye tracking data were exported through the Tobii Pro Lab and organized for each test condition. Among the exported measures, we analyzed several key eye-tracking metrics: fixation durations, saccade durations, number of fixations, number of saccades, saccade amplitude, saccade velocity, and pupil diameter. These metrics were chosen for their established relationships with visual attention, emotional arousal, and cognitive workload in humans (Skaramagkas et al., 2023). For more detailed description regarding the cursor and eye movement measures, please refer to Appendix A or the corresponding references.

For all collected and calculated measures, we verified data normality using Ryan-Joiner tests, which confirmed that most measures were either normally distributed or showed no serious deviations from normal distribution. Subsequently, we conducted repeated measures analysis of variance (RM-ANOVA) at a significance level of 0.05. To address the multiple comparison problem inherent in multiway ANOVA, we applied the Benjamini-Hochberg procedure to control the false discovery rate (Benjamini &

Hochberg, 1995), and reported adjusted p-values throughout our analyses. It is noteworthy to mention that both $W$ and $A$ were incorporated into the ANOVA model for analysis. Additionally, for the factors with three levels (i.e., $W$ and $A$), we assessed the assumption of sphericity and applied the Greenhouse-Geisser correction where necessary. All data processing and statistical analyses were performed using Python 3.10.11 and R 4.4.1.

## 4. Results

For clarity, we focus on key findings related to four task factors. Full RM-ANOVA tables of all dependent variables are provided in the Supplementary Materials.

### 4.1. Performance

The main effects analysis revealed several significant findings (see Table 3 and Figure 4). For *MT*, both *EH* and *TS* showed significant effects. *ER* was significantly affected by *EH*. *TP* demonstrated significant main effects for *HH* and *TN*.

In addition, we found a few significant two-way interaction effects for $EH \times HH$ ($F_{1,34} = 4.719, p = 0.037, \eta_p^2 = 0.122$) and $TN \times TS$ ($F_{1,34} = 4.315, p = 0.045, \eta_p^2 = 0.113$) on *MT*, and $TN \times TS$ ($F_{1,34} = 5.32, p = 0.027, \eta_p^2 = 0.135$) on *TP* (Figure 5). However, upon examining the graphical representations and associated values, no notable patterns emerged that warranted further discussion. No significant interaction effects were found on *ER*. Additionally, no significant three-way or four-way interaction effects among four task factors were observed in our analysis.

**Table 3.** $F$, $p$, and $\eta_p^2$ values from the RM-ANOVA results showing the effects of four task factors on performance metrics. P-values are adjusted for false discovery rate using the Benjamini–Hochberg procedure. *$p< 0.05$, **$p< 0.01$, ***$p< 0.001$.

| Performance | MT | | | ER | | | TP | | |
|---|---|---|---|---|---|---|---|---|---|
| | $F$ | $p$ | $\eta_p^2$ | $F$ | $p$ | $\eta_p^2$ | $F$ | $p$ | $\eta_p^2$ |
| **EH** | 52.131 | ***<0.001 | 0.585 | 16.933 | **0.004 | 0.314 | 0.285 | 0.858 | 0.008 |
| **HH** | 0.207 | 0.821 | 0.006 | 5.534 | 0.217 | 0.130 | 17.722 | **0.002 | 0.324 |
| **TS** | 96.063 | ***<0.001 | 0.722 | 1.201 | 0.721 | 0.031 | 3.691 | 0.271 | 0.091 |
| **TN** | 0.183 | 0.823 | 0.005 | 0.018 | 0.940 | 0.000 | 33.411 | ***<0.001 | 0.475 |

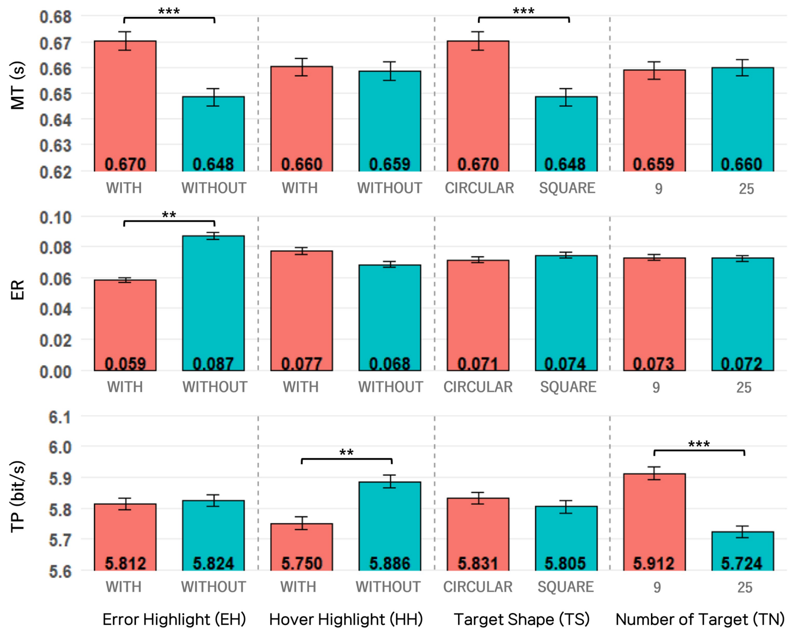



**Figure 4.** Mean performance measures by each task factor. Error bars indicate 95% confidence intervals. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. The y-axis do not start at 0 for MT and TP.

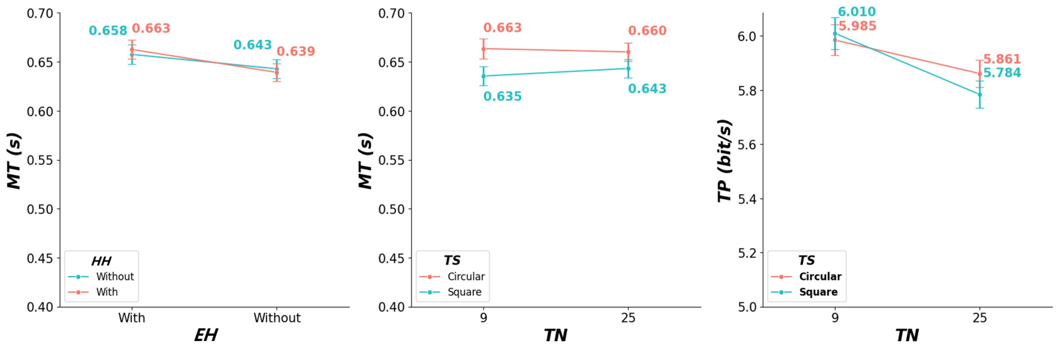



**Figure 5.** Mean performance measures with significant interaction effects between four task factors. Error bars indicate 95% confidence intervals.

### *4.2. Behavior*

#### *4.2.1. Cursor Movement*

Analysis of cursor movement measures revealed several significant effects (see Table 4 and Figure 6(a)). Target re-entry was more frequent under the condition of $EH_{with}$, while the $HH$ condition showed significantly higher counts under $HH_{without}$. Additionally, the number of target re-entries significantly increased under the condition of $TN_{25}$. Task axis crossing significantly increased under $HH_{without}$ and was significantly more frequent with $TS_{circular}$ compared to $TS_{square}$. Movement direction showed significant differences across the factors $EH$, $TS$, and $TN$, with significantly more changes observed under the conditions of $EH_{with}$, $TS_{circular}$, and $TN_{25}$, respectively. In the orthogonal direction, significantly more changes were observed under both $TS_{circular}$ and $HH_{without}$. Movement variability was significantly influenced only by $TN$, showing significantly higher values under $TN_9$. Errors in movement also showed significant differences depending on the conditions. Significantly more movement errors occurred under the conditions of $EH_{without}$, $TS_{square}$, or $TN_9$. Finally, movement offset was significantly influenced only by $TN$, with significantly lower values observed under $TN_{25}$.

**Table 4.** $F$, $p$, and $\eta_p^2$ values from the RM-ANOVA results showing the effects of four task factors on cursor movement metrics. P-values are adjusted for false discovery rate using the Benjamini–Hochberg procedure. *$p< 0.05$, **$p< 0.01$, ***$p< 0.001$.

| **Cursor** | **EH** | | | **HH** | | | **TS** | | | **TN** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Movement** | $F$ | $p$ | $\eta_p^2$ | $F$ | $p$ | $\eta_p^2$ | $F$ | $p$ | $\eta_p^2$ | $F$ | $p$ | $\eta_p^2$ |
| **Target Re-entry** | 8.716 | 0.057 | 0.191 | 41.950 | ***<0.001 | 0.531 | 1.199 | 0.707 | 0.031 | 15.633 | **0.004 | 0.297 |
| **Task Axis Crossing** | 2.098 | 0.546 | 0.054 | 28.319 | ***<0.001 | 0.434 | 20.705 | ***<0.001 | 0.359 | 3.156 | 0.335 | 0.079 |
| **Movement Direction Change** | 36.427 | ***<0.001 | 0.496 | 2.536 | 0.379 | 0.064 | 31.248 | ***<0.001 | 0.458 | 37.486 | ***<0.001 | 0.503 |
| **Orthogonal Direction Change** | 2.569 | 0.711 | 0.065 | 27.710 | ***<0.001 | 0.428 | 15.762 | **0.004 | 0.299 | 5.770 | 0.193 | 0.135 |
| **Movement Variability** | 3.426 | 0.379 | 0.085 | 4.510 | 0.255 | 0.109 | 4.941 | 0.227 | 0.118 | 12.713 | *0.016 | 0.256 |
| **Movement Error** | 10.834 | *0.023 | 0.226 | 2.079 | 0.603 | 0.053 | 25.992 | ***<0.001 | 0.413 | 12.535 | *0.014 | 0.253 |
| **Movement Offset** | 0.085 | 0.929 | 0.002 | 2.027 | 0.540 | 0.052 | 0.058 | 0.929 | 0.002 | 58.928 | ***<0.001 | 0.614 |

### *4.2.2. Eye Movement*

Analysis of the eye movement measures revealed several significant effects (see Table 5 and Figure 6(b)). Fixation duration showed significant differences depending on the factors, with significantly longer durations observed under the conditions of $EH_{with}$ and $TS_{circular}$ separately. The number of fixations was significantly lower under the conditions of $TN_{25}$. The saccade amplitude and saccade velocity showed significant differences only under the condition of $TN$, with values being significantly higher at $TN_{25}$. Among the four factors, only $HH$ had a significant effect on pupil diameter, with significantly higher values observed under the condition $HH_{without}$ compared to $HH_{with}$. Lastly, in saccade duration and the number of saccade, no statistically significant differences were found across the task factors.

**Table 5.** $F$, $p$, and $\eta_p^2$ values from the RM-ANOVA results showing the effects of four task factors on eye movement metrics. P-values are adjusted for false discovery rate using the Benjamini–Hochberg procedure. *$p<$ 0.05, **$p<$ 0.01, ***$p<$ 0.001.

| Eye | EH | | | HH | | | TS | | | TN | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Movement | $F$ | $p$ | $\eta_p^2$ | $F$ | $p$ | $\eta_p^2$ | $F$ | $p$ | $\eta_p^2$ | $F$ | $p$ | $\eta_p^2$ |
| Fixation Duration | 37.777 | ***<0.001 | 0.505 | 0.266 | 0.812 | 0.007 | 46.214 | ***<0.001 | 0.555 | 0.004 | 0.967 | 0.000 |
| Saccade Duration | 1.224 | 0.800 | 0.032 | 0.125 | 0.900 | 0.003 | 0.998 | 0.816 | 0.026 | 6.078 | 0.232 | 0.141 |
| Number of Fixation | 8.125 | 0.075 | 0.180 | 0.367 | 0.882 | 0.010 | 1.716 | 0.587 | 0.044 | 22.347 | ***<0.001 | 0.377 |
| Number of Saccade | 1.757 | 0.763 | 0.045 | 0.006 | 0.970 | <0.001 | 0.584 | 0.776 | 0.016 | 0.766 | 0.763 | 0.020 |
| Saccade Amplitude | 2.735 | 0.354 | 0.069 | 0.088 | 0.910 | 0.002 | 3.677 | 0.274 | 0.090 | 88.871 | ***<0.001 | 0.706 |
| Saccade Velocity | 0.210 | 0.949 | 0.006 | 0.032 | 0.949 | 0.001 | 0.139 | 0.949 | 0.004 | 11.179 | *0.024 | 0.232 |
| Pupil Diameter | 0.032 | 0.925 | 0.001 | 27.612 | ***<0.001 | 0.427 | 1.107 | 0.899 | 0.029 | 0.225 | 0.925 | 0.006 |

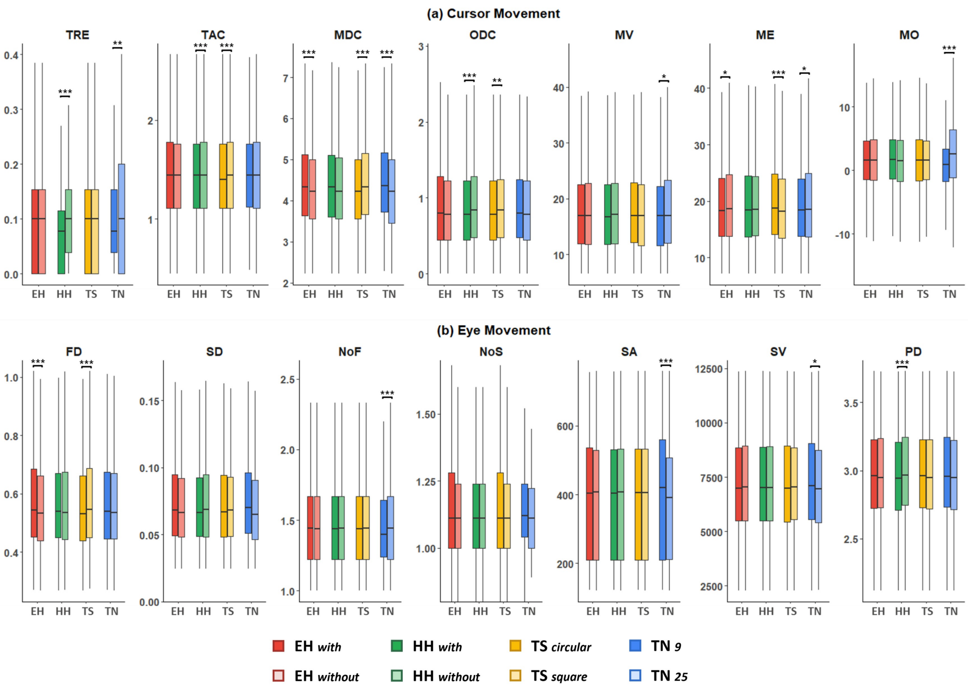



**Figure 6.** Mean behavior measures: (a) cursor movement and (b) eye movement by each task factor. Error bars indicate 95% confidence intervals. TRE: Target re-entry, TAC: Target axis crossing, MDC: Movement direction change, ODC: Orthogonal direction change, MV: Movement variability, ME: Movement error, MO: Movement offset, FD: Fixation duration, SD: Saccade duration, NoF: Number of fixations, NoS: Number of saccades, SA: Saccade amplitude, SV: Saccade velocity, PD: Pupil diameter. *$p<$ 0.05, **$p<$ 0.01, ***$p<$ 0.001.

### *4.3. Subjective Feedback*

Figure 7 shows the percentage of participants felt influenced by each task factor. Regarding *EH*, 32 participants noted its impact on their performance, while 3 reported no difference. The majority stated that *EH* increased their caution and focus, often leading to slower movements to ensure accuracy, which aligns with our previous findings on *MT*. Conversely, without *EH*, participants frequently overlooked errors or felt less urgency. For *HH*, 33 participants indicated its influence on their performance, while 2 observed no difference. Most noted that the visual feedback provided clear confirmation, enabling faster and more confident movements. In its absence, participants reported increased caution and precision. Concerning *TS*, 30 participants indicated that it affected their performance, while 5 reported no difference. Of those who noticed an effect, 14 preferred circular targets, 10 found square targets easier to use, and 6 mentioned that their preference varied based on specific circumstances. With respect to *TN*, 27 participants reported that the number of targets influenced their experience, while 8 noticed no significant difference. Most found fewer targets made the task easier and less stressful, although some mentioned that more targets helped maintain focus. A few observed that when targets were fewer and more spread out, they were more difficult to select.

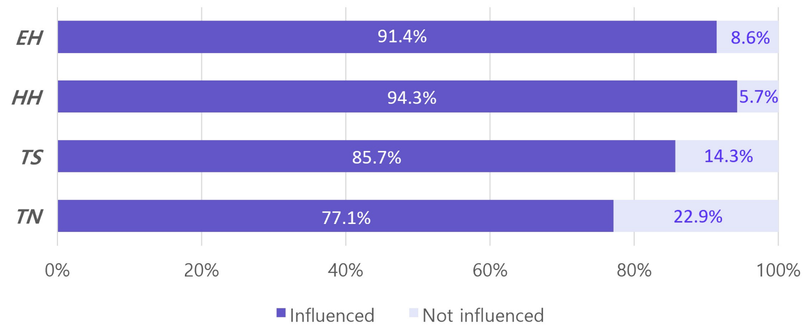



**Figure 7.** Percentage of participants felt influenced by each task factor.

## 5. Discussion

Fitts' law is regarded as one of the most robust and well-established models in HCI, particularly in the context of desktop GUIs (MacKenzie & Buxton, 1992; Soukoreff & MacKenzie, 2004; Zhai, Kong, & Ren, 2004). Therefore, our conclusions are based on the assumption that better task performance can be achieved by minimizing nuisance factors that hinder task execution. We believe this indicates more general and natural human behavior could be emerged, enabling a more accurate measurement of task performance. In other words, it might suggest that the tasks were well-designed to better capture the essence of the pointing behavior Fitts' law describes.

Grounded in this assumption, we proceed to examine how each of the four task factors influenced user performance, along with potential underlying reasons, drawing on participants' behavioral measures and subjective feedback in the following sections.

### *5.1. Error Highlight*

In evaluating user performance, we found that $EH_{with}$, compared to $EH_{without}$, led to a significant increase in $MT$ between targets and a significant decrease in $ER$ for selection. The increase in $MT$ can be explained through the analysis of participants' gaze fixation. Results indicated that *EH* significantly affected fixation duration. This suggests that visual stimuli from error feedback may have increased fixation, as participants' gaze was directed not only to the targets but also to the *EH*. The increase in fixation duration likely occurred because participants, aware of their mistakes, spent more time carefully focusing on the task to improve their performance (Negi & Mitra, 2020; Skaramagkas et al., 2023). Figure 8 supports this observation by illustrating the increased fixation duration in the presence of $EH_{with}$ compared to $EH_{without}$.

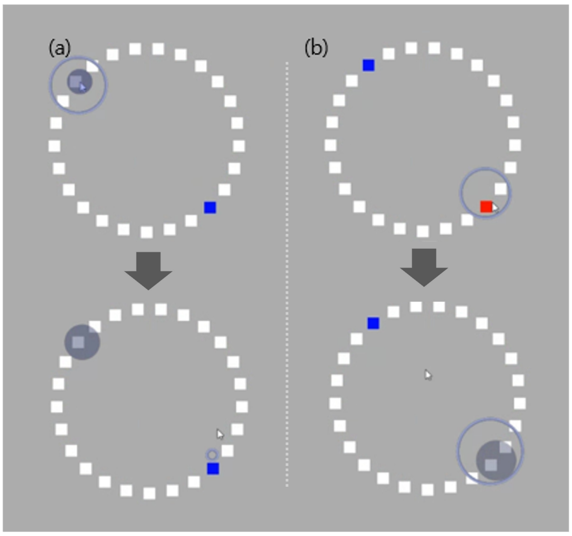



**Figure 8.** The difference in fixation duration and location according to *EH* condition. The circles represent the points where eye fixation occurred. In (a), under the condition of without error highlight, the fixation shifts to the next target simultaneously as the movement toward the next target begins. However, in (b), under the condition of with error highlight, even though the movement is directed toward the next target, the fixation remains at the feedback of the error.

The conditions for providing feedback have been demonstrated in VR-based research, and studies showing that auditory error feedback has a significant effect on error reduction are consistent with the findings of this study (Batmaz & Stuerzlinger, 2021; McAnally & Wallis, 2024). The decrease in error rate is, as expected, a natural consequence of improved participant attention due to error feedback. Upon receiving feedback, participants exhibited behavioral adjustments aimed at reducing errors. Participants' behavior related to errors could be directly analyzed through cursor movements via mouse. The significant increase in target re-entry when *EH* was provided could be interpreted as a corrective movement in response to feedback stimuli. However, it might also result from hand tremors caused by lack of precise control, arising from the speed-accuracy trade-off in performing the task rapidly (MacKenzie & Isokoski, 2008). This can be related to the significance observed in cursor movement direction changes and movement errors as well. Furthermore, through movement direction changes and movement errors, we can confirm that participants' cursor movements significantly deviated from straight-line trajectories. This aligns with the observed increase in *MT*. These findings suggest that the psychological discomfort of participants due to error feedback may have influenced the spatial characteristics of cursor movements (i.e., direction changes) (Yamauchi, 2013).

In conclusion, providing *EH* reduces *ER* but simultaneously increases *MT*. Given that there's no significant effect on *TP*, it appears that *EH* does not necessarily improve or decrease overall performance, but rather influences participants to shift their strategy towards prioritizing accuracy over speed. Having a high error rate is generally not ideal in a Fitts' Law study, as it indicates that participants are making frequent mistakes, which can skew the results. Therefore, providing *EH* can be recommended in most typical situations. However, in specific scenarios where the majority of participants are overly conscious about accuracy and not performing "as quickly as possible", removing *EH* could potentially prompt a strategic adjustment in their approach.

### *5.2. Hover Highlight*

Hovering, which notifies users when the cursor enters the target, serves as a form of task performance support. Task performance with *HH* resulted in lower *TP*, whereas significantly higher *TP* was observed in task conditions without *HH*. These results suggest that *HH* had a negative impact on participants' task performance. Target re-entry significantly decreased when *HH* was provided, which can be attributed to participants ceasing corrective cursor movements after initially entering the target in response to the received feedback. Figure 9 shows that it is highly possible that participants either didn't realize they had not selected the target or incorrectly believed they had made a correct selection, due to the *HH* they had already received.

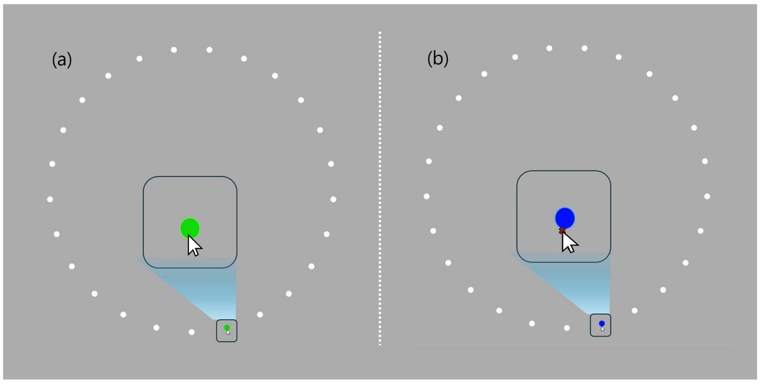



**Figure 9.** The difference in cursor location between when *HH* appears and the actual selection point. In (a), the participant recognizes the green indicator, and in (b), the participant presses the target after recognizing it. $HH_{with}$ gives the participant excessive confidence, which, in this case, reduces the participant's effort to re-enter the target with the mouse, as shown by the decrease in corrective actions.

As pupil diameter is associated with the user's mental effort and is proportional to task difficulty (S. Chen, Epps, Ruiz, & Chen, 2011; Ehlers, Strauch, & Huckauf, 2018), the decreased pupil diameter when *HH* was provided implies that participants may subjectively perceive *HH* as useful. The analysis of cursor movement showed similar results. In tasks without *HH*, task axis crossing and orthogonal direction change significantly increased. Due to the characteristics of *HH* being perceived similar to feedback for task success, participants might have felt the task was more difficult when the indicator was not present, which could have resulted in increased cursor movement changes. In participant interviews, many stated that target selection was easier and more comfortable when $HH$ was provided. One participant (P07) even mentioned that the color change that appeared as an indication helped improve concentration on the task.

Despite participants' subjective perception of reduced difficulty, the fact that more accurate performance measurements are possible without $HH$ suggests that providing $HH$ in the multi-directional tapping task may not be appropriate.

### *5.3. Target Shape*

$TS$ showed a significant effect only on $MT$ in performance measures. Previous studies have addressed $TS$ compared to other factors (Lin & Cheng, 2022; Sheikh & Hoffmann, 1994). Although they did not use the multi-directional tapping tasks, they sought to confirm significant differences in $MT$ when selecting targets of different shapes. This study aligns with previous research in that it showed significant differences in $MT$ depending on the $TS$. In this study, there was a significant decrease in $MT$ when square targets were provided. As shown in Figure 10, despite having the same width and height, square targets have a larger area compared to circular targets, allowing users to reach the target with less movement distance. Lin et al. (Lin & Cheng, 2022) stated that the area of geometric targets affects perceived difficulty, specifically, perceived difficulty is influenced by both area and central width. Through experiments, it was explained that when a circular target and a diamond-shaped target have the same central width, users perceive the diamond-shaped target as more difficult than the circular target because the circular target has a larger area. This principle can be applied equally to circles and squares.

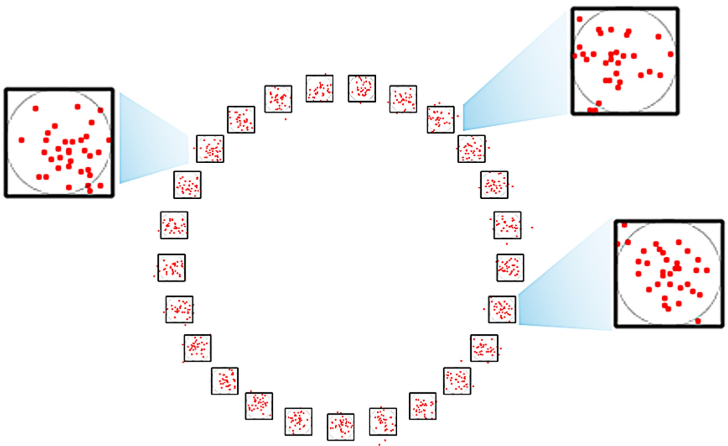


**Figure 10.** The distribution of selection points (in red) for square targets compared to circulars with the same width. The square represents the actual target size, while the circular illustrates the selection space for targets of equivalent width, showing that participants use more space when interacting with square targets.

Task axis crossing and movement direction change are related metrics (MacKenzie et al., 2001). In this experiment, we confirmed through the results of these two metrics that the inaccuracy of cursor movement decreased when using square targets. Integrating this with the earlier results on $MT$, we can conclude that when targets were provided as squares rather than circles, users' movements were faster.

According to the gaze data, fixation duration increased for circular targets compared to square targets. This result can be interpreted as increased attention being required to recognize circular targets (Shojaeizadeh, Djamasbi, & Trapp, 2016). In a post-experiment interview, one participant (P17) stated, *"Circles are more comfortable to look at, but squares were easier to click"*. As shown in the previous results, while users may perceive circular targets as visually more comfortable and easier, the verified results demonstrate that selecting square targets is actually more effective, very likely due to the larger area size.

In conclusion, it is highly likely that the difference in target area determined by the target shape had a greater impact than the geometric characteristics of the shape itself. Given that when the target width is the same for both circular targets (diameter) and square targets (side length), the square target has a larger area ($W^2(1 - \frac{\pi}{4})$ wider to be exact), leading to better performance, it would be incorrect to simply consider square targets as superior. We believe that the task protocol should explicitly state that the area must be taken into account when determining target width for various target shapes. However, additional experiments are necessary to draw more definitive conclusions. These should include scenarios where shapes differ but have equal areas, and task configurations where targets are rotated and aligned with the direction of task axis.

### *5.4. Number of Targets*

When 25 targets were arranged, the spacing was narrower compared to 9 targets, so the degree of overlap between targets was much higher (Blanch & Ortega, 2011). This might lead to mistakes due to afterimages cause by focused attention in tasks requiring quick and accurate movements. The 9 targets, with their wider spacing, could allow attention to spread, weakening afterimages compared to the 25 targets (Baijal & Srinivasan, 2009). This can also be explained by the increased number of target re-entries; the increased difficulty of perceiving each of 25 targets likely caused the mouse cursor to enter and exit targets more frequently. Additionally, the significant difference in movement direction change showed that when moving between 25 targets, the cursor could not move in a straight line and experienced more left-right oscillation around the axis connecting the targets, confirming that the number of targets significantly affected participants' behavior.

In terms of task efficiency, $TP$ indicated improved participant performance with fewer targets. This finding can be elucidated through the saccade data (see Figure 11), which represents participants' eye movement behavior. While eye movement data serves as a useful indicator of user behavior, it is important to consider environmental and usage contexts (Hyrskykari, Majaranta, & Räihä, 2003; Majaranta, Räihä, Hyrskykari, & Špakov, 2019). Saccades, representing momentary gaze strokes, showed significant effects in amplitude and velocity.

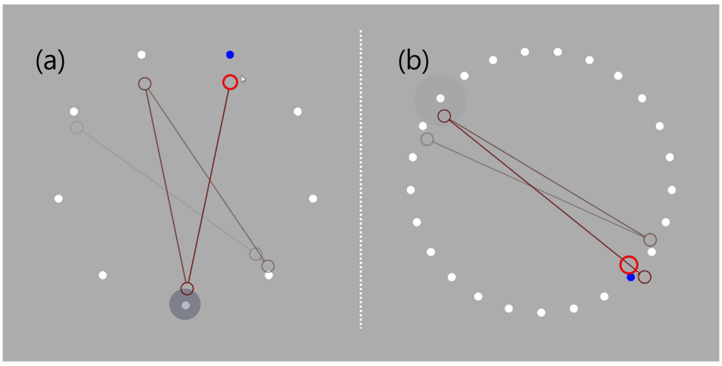



**Figure 11.** The difference in saccade amplitude between $TN_9$ and $TN_{25}$ caused by the distance from the fixation point to the target. In this figure, the red lines and red circles represent the saccade and fixation paths over a 3-second period. In (a), when there are 9 targets, the fixation position is generally far from the target, while in (b), when there are 25 targets, the fixation position is generally much closer to the target than when there are 9 targets.

These results suggest that the number of targets functions as a proxy for search volume. As the number of targets increases, so does the amount of visual searching required, leading to increased cognitive workload (Ball & Richardson, 2022). This interpretation is corroborated by participant interviews, where many reported that tasks with fewer targets felt easier and more comfortable. The coherence between the quantitative saccade data and qualitative participant feedback provides a robust foundation for understanding the impact of target number on task performance and user experience.

It is easy to assume that performing tasks with a larger number of targets generally leads to a more reliable experimental design, which is partly true due to the increased number of trials and testing of more diverse movement directions. However, even with fewer targets, the issue of limited trials can be addressed by increasing the number of repetitions (rounds of trials), and diverse movement directions can be incorporated by slightly varying the orientation of the circular layout. In fact, having fewer targets might be advantageous as it allows testing of wider angle variations in the task axis, whereas a higher number of targets results in narrower angle changes due to the zigzag nature of movements in the multi-directional tapping task. Most importantly, from the perspective of making participants comfortable, increasing satisfaction, and improving task performance (throughput), having fewer targets seems to have no drawbacks.

It is important to note that these findings are based on a 24-inch monitor, where limited spacing between targets can amplify density effects. On larger displays or in immersive VR, longer target amplitudes could reduce visual crowding, making density less influential and potentially altering the optimal number of targets. Future work should examine whether the advantages of fewer targets hold across such platforms. Of course, this does not mean that fewer is always better (e.g., would 3 targets be better than 9?). While the optimal number could be determined through follow-up studies, at this stage we can conclude that there seems to be little merit in considering a very high number of targets as seen in the ISO/TS 9241-411 (Figure 2) or some related works (Table 2).

### 5.5. Limitations and Future Work

This study has several limitations. The limited levels of task factors examined may not fully capture the complexity of human-task interaction. Expanding the range of variable levels could provide a more nuanced understanding of factors and reveal important variations in user performance and behavior. In addition, the experiment was only tested in desktop GUI using a mouse, limiting the generalizability to other types of input devices and UIs. In particular, it remains unclear whether these four task factors would have the same impact in mobile UIs using touch interaction or in spatial UIs using VR controllers or hand gestures. Hence, future research can focus on two key areas:

- Expanding the granularity of each factor's levels to gain a more nuanced understanding of their influence on Fitts' law performance. This could involve exploring a wider spectrum of target sizes, shapes, numbers, feedback mechanisms, or even other potential factors such as indication of the subsequent target and direction of the target sequence (clockwise/anti-clockwise).
- Extending this research to other computing environments could be valuable. In particular, virtual and augmented reality interfaces are not only becoming more prevalent but also introduce additional task factors such as target depth, target visual angle, stereoscopic vision, and spatial orientation, making them worthy of further investigation. Future research could serve as an initial step towards establishing a 3D Fitts' law task protocol.

We also acknowledge that the task protocol, which reflects the task factors examined in this study, may reduce external validity and represent real-world usage less accurately. However, we would argue that all four investigated task factors deal with very fundamental elements and do not involve a level of detail that could significantly distance them from representation of real-world scenarios.

## 6. Conclusion

This study investigated the effects of four task factors: Error Highlight, Hover Highlight, Target shape, and Number of Targets, on user performance and behavior in multi-directional tapping tasks. Our findings reveal that these factors had substantial effects on user performance and behavior.

### 6.1. Implications

For researchers planning to employ multi-directional tapping tasks in future studies, our findings provide several important considerations regarding the task configuration:

- **Error Highlight:** While *EH* is generally effective in sustaining low error rates, its applications should be approached with caution with contexts where task speed is prioritized, as its removal may prompt participants to adopt a faster yet potentially less accurate strategy.
- **Hover Highlight:** Although *HH* is often perceived by users as beneficial, empirical findings indicate that it can inadvertently impair performance, potentially by fostering premature task completion or reducing error correction due to misinterpretation as confirmation of target selection.
- **Target Shape:** Square targets demonstrated superior performance in terms of

movement time compared to circular targets of identical width, likely due to their larger effective area. However, to ensure fair and valid comparisons in pointing tasks, target area, in addition to width, should be explicitly controlled.

- **Number of Targets:** Comparable experimental validity can be achieved with a reduced number of targets when accompanied by increased repetitions and varied layouts, which may also enhance user comfort, satisfaction, and performance, thereby limiting the necessity of employing very high target counts commonly observed in standards and prior studies.

These results underscore the importance of carefully considering these factors in the design of pointing tasks and suggest a need for more standardized protocols in multi-directional tapping task experiments. Our findings provide valuable implications for researchers in human-computer interaction, particularly those evaluating input devices and interaction techniques. This study represents a significant step towards establishing more consistent and comparable methodologies in Fitts' law research and pointing task evaluations.

## CRediT authorship contribution statement

**Yuhwa Hong**: Conceptualization, Methodology, Writing – original draft, Investigation, Visualization. **Haejun Kim**: Methodology, Writing – original draft, Software, Data curation, Formal analysis. **Jihae Yu**: Methodology, Investigation. **Heedo Shin**: Methodology, Investigation. **Xiaoqun Yu**: Methodology, Validation. **Shuping Xiong**: Project administration, Supervision, Writing – review & editing. **Woojoo Kim**: Conceptualization, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing, Funding acquisition.

## Acknowledgement

This work was supported by Innovative Human Resource Development for Local Intellectualization Program through the Institute of Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT)(IITP-2025-RS-2023-00260267) and the National Research Foundation of Korea(NRF) grants funded by the Korea government(MSIT)(RS-2024-00343882, RS-2023-00242528).

## Declaration of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

## Data availability statement

The data that support the findings of this study are available from the corresponding author, Woojoo Kim, upon reasonable request.

## References


Amini, M., Stuerzlinger, W., Teather, R. J., & Batmaz, A. U. (2025). A systematic review of fitts' law in 3d extended reality. In *Proceedings of the 2025 chi conference on human factors in computing systems.* New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3706598.3713623 doi:

Avsar, H., Fischer, J. E., & Rodden, T. (2016, 9). Future flight decks: Impact of +gz on touchscreen usability. *Proceedings of the International Conference on Human-Computer Interaction in Aerospace, HCI-Aero 2016*. Retrieved from https://dl.acm.org/doi/10.1145/2950112.2964592 doi:

Baijal, S., & Srinivasan, N. (2009, 12). Types of attention matter for awareness: A study with color afterimages. *Consciousness and Cognition*, *18*, 1039-1048. doi:

Ball, L. J., & Richardson, B. H. (2022). Eye movement in user experience and human–computer interaction research. *Neuromethods*, *183*, 165-183. Retrieved from https://link.springer.com/protocol/10.1007/978-1-0716-2391-6_10 doi:

Bateman, S., Mandryk, R. L., Gutwin, C., & Xiao, R. (2013, 5). Analysis and comparison of target assistance techniques for relative ray-cast pointing. *International Journal of Human-Computer Studies*, *71*, 511-532. doi:

Batmaz, A. U., & Stuerzlinger, W. (2021, 3). Effects of different auditory feedback frequencies in virtual reality 3d pointing tasks. *Proceedings - 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops, VRW 2021*, 189-194. doi:

Batmaz, A. U., & Stuerzlinger, W. (2022). Effective throughput analysis of different task execution strategies for mid-air fitts' tasks in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, *28*(11), 3939–3947.

Batmaz, A. U., Yu, K., Liang, H. N., & Stuerzlinger, W. (2022, 12). Improving effective throughput performance using auditory feedback in virtual reality. *Proceedings - SUI 2022: ACM Conference on Spatial User Interaction.* Retrieved from https://dl.acm.org/doi/10.1145/3565970.3567702 doi:

Beelders, T. R., & Blignaut, P. J. (2012). Using eye gaze and speech to simulate a pointing device. *Eye Tracking Research and Applications Symposium (ETRA)*, 349-352. Retrieved from https://dl-acm-org.libra.kaist.ac.kr/doi/10.1145/2168556.2168634 doi:

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289–300.

Bernardos, A. M., Gómez, D., & Casar, J. R. (2016, 4). A comparison of head pose and deictic pointing interaction methods for smart environments. *International Journal of Human-Computer Interaction*, *32*, 325-351. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/10447318.2016.1142054 doi:

Blanch, R., & Ortega, M. (2011). Benchmarking pointing techniques with distractors: adding a density factor to fitts' pointing paradigm. In *Proceedings of the sigchi conference on human factors in computing systems* (p. 1629–1638). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/1978942.1979180 doi:

Boritz, J., Booth, K., human performance, W. C., & undefined 1991. (1991). Fitts's law studies of directional mouse movement. *graphicsinterface.orgJ Boritz, KS Booth, WB Cowanhuman performance, 1991•graphicsinterface.org*. Retrieved from http://graphicsinterface.org/wp-content/uploads/gi1991-28.pdf

Browning, G., & Teather, R. J. (2014). Screen scaling: Effects of screen scale on moving target selection. *Conference on Human Factors in Computing Systems - Proceedings*, 2053-2058. Retrieved from https://dl.acm.org/doi/10.1145/2559206.2581227 doi:

Cassidy, B., Read, J. C., & MacKenzie, I. S. (2019). Fittsfarm: Comparing children's drag-and-drop performance using finger and stylus input on tablets. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelli-*

*gence and Lecture Notes in Bioinformatics)*, *11748 LNCS*, 656-668. Retrieved from https://link.springer.com/chapter/10.1007/978-3-030-29387-1_38 doi:

Chen, H. J., Lin, C. J., & Lin, P. H. (2019, 7). Effects of control-display gain and postural control method on distal pointing performance. *International Journal of Industrial Ergonomics*, *72*, 45-53. doi:

Chen, S., Epps, J., Ruiz, N., & Chen, F. (2011). Eye activity as a measure of human mental effort in hci. *International Conference on Intelligent User Interfaces, Proceedings IUI*, 315-318. Retrieved from https://dl.acm.org/doi/10.1145/1943403.1943454 doi:

Cheong, K. K., Kim, I., Park, S. K., & Park, Y. J. (2011, 8). User performance measures for evaluating interactive tv pointing devices. *IEEE Transactions on Consumer Electronics*, *57*, 1236-1244. doi:

Cockburn, A., Ahlström, D., & Gutwin, C. (2012, 3). Understanding performance in touch selections: Tap, drag and radial pointing drag with finger, stylus and mouse. *International Journal of Human-Computer Studies*, *70*, 218-233. doi:

Cockburn, A., & Brock, P. (2006). Human on-line response to visual and motor target expansion. Retrieved from http://ir.canterbury.ac.nz/bitstream/10092/90/1/12602894_Main.pdf

Cuaresma, J., & Mackenzie, I. S. (2017). Fittsface: Exploring navigation and selection methods for facial tracking. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10278 LNCS*, 403-416. Retrieved from https://link.springer.com/chapter/10.1007/978-3-319-58703-5_30 doi:

Draghici, O., Batkin, I., Bolic, M., & Chapman, I. (2014). Performance evaluation of the mouthpad. *IEEE MeMeA 2014 - IEEE International Symposium on Medical Measurements and Applications, Proceedings*. doi:

Dube, T. J., Ren, Y., Limerick, H., Mackenzie, I. S., & Arif, A. S. (2022, 11). Push, tap, dwell, and pinch: Evaluation of four mid-air selection methods augmented with ultrasonic haptic feedback. *Proceedings of the ACM on Human-Computer Interaction*, *6*, 207-225. Retrieved from https://dl.acm.org/doi/10.1145/3567718 doi:

Ehlers, J., Strauch, C., & Huckauf, A. (2018, 11). A view to a click: Pupil size changes as input command in eyes-only human-computer interaction. *International Journal of Human-Computer Studies*, *119*, 28-34. doi:

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*, 381-391. doi:

Fogg, B. J., & Nass, C. (1997, 5). Silicon sycophants: the effects of computers that flatter. *International Journal of Human-Computer Studies*, *46*, 551-561. doi:

Gillan, D. J., Holden, K., Adam, S., Rudisill, M., & Magee, L. (1990, 3). How does fitts' law fit pointing and dragging? *Conference on Human Factors in Computing Systems - Proceedings*, 227-234. Retrieved from https://dl.acm.org/doi/10.1145/97243.97278 doi:

Gori, J., & Bellut, Q. (2023, 4). Positional variance profiles (pvps): A new take on the speed-accuracy trade-off. *Conference on Human Factors in Computing Systems - Proceedings*, 16. Retrieved from https://dl.acm.org/doi/10.1145/3544548.3581071 doi:

Herring, S. R., Trejo, A. E., & Hallbeck, M. S. (2010, 1). Evaluation of four cursor control devices during a target acquisition task for laparoscopic tool control. *Applied Ergonomics*, *41*, 47-57. doi:

Hertzum, M., & Hornbæk, K. (2013, 4). The effect of target precuing on pointing with mouse and touchpad. *International Journal of Human-Computer Interaction*, *29*, 338-350. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/10447318.2012.711704 doi:

Hoffmann, E., & Sheikh, I. H. (1994). Effect of varying target height in a fitts' movement task. *Ergonomics*, *37*, 1071-1088. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/00140139408963719 doi:

Hyrskykari, A., Majaranta, P., & Räihä, K.-J. (2003). Proactive response to eye movements.

*Human-Computer Interaction—INTERACT 2003*.

ISO. (2000). *Iso 9241-9. ergonomic requirements for office work with visual display terminals (vdts) — part 9: Requirements for non-keyboard input devices.* International Organization for Standardization.

ISO. (2012). *Iso/ts 9241-411. international standard: ergonomics of human-system interaction — part 411: Evaluation methods for the design of physical input devices.* International Organization for Standardization.

Janthanasub, V., & Meesad, P. (2015). Evaluation of a low-cost eye tracking system for computer input. *Applied Science and Engineering Progress*, *8*(3), 185–196.

José, M. A., & Lopes, R. D. D. (2015, 1). Human-computer interface controlled by the lip. *IEEE Journal of Biomedical and Health Informatics*, *19*, 302-308. doi:

Kasahara, N., Oba, Y., Yamanaka, S., Stuerzlinger, W., & Miyashita, H. (2023). Throughput and effective parameters in crossing. In *Extended abstracts of the 2023 chi conference on human factors in computing systems.* New York, NY, USA: Association for Computing Machinery. Retrieved from `https://doi.org/10.1145/3544549.3585817` doi:

Katsuragawa, K., Pietroszek, K., Wallace, J. R., & Lank, E. (2016, 6). Watchpoint: Freehand pointing with a smartwatch in a ubiquitous display environment. *Proceedings of the Workshop on Advanced Visual Interfaces AVI*, *07-10-June-2016*, 128-135. Retrieved from `https://dl.acm.org/doi/10.1145/2909132.2909263` doi:

Kim, S., Lee, B., Gemert, T. V., & Oulasvirta, A. (2020, 4). Optimal sensor position for a computer mouse. *Conference on Human Factors in Computing Systems - Proceedings*. Retrieved from `https://dl.acm.org/doi/10.1145/3313831.3376735` doi:

Kim, W., & Xiong, S. (2022, 5). Viewfindervr: configurable viewfinder for selection of distant objects in vr. *Virtual Reality*. Retrieved from `https://link.springer.com/10.1007/s10055-022-00649-z` doi:

Ko, Y. J., Zhao, H., Kim, Y., Ramakrishnan, I. V., Zhai, S., & Bi, X. (2020, 10). Modeling two dimensional touch pointing. *UIST 2020 - Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, 858-868. Retrieved from `https://dl.acm.org/doi/10.1145/3379337.3415871` doi:

Kouroupetroglou, G., Pino, A., Balmpakakis, A., Chalastanis, D., Golematis, V., Ioannou, N., & Koutsoumpas, I. (2012). Using wiimote for 2d and 3d pointing tasks: Gesture performance evaluation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7206 LNAI*, 13-23. Retrieved from `https://link.springer.com/chapter/10.1007/978-3-642-34182-3_2` doi:

Li, Y., Liu, J., Huang, J., Zhang, Y., Peng, X., Bian, Y., & Tian, F. (2024, 11). Evaluating the effects of user motion and viewing mode on target selection in augmented reality. *International Journal of Human-Computer Studies*, *191*, 103327. doi:

Li, Y., Sarcar, S., Kim, K., Tu, H., & Ren, X. (2022, 9). Designing successive target selection in virtual reality via penetrating the intangible interface with handheld controllers. *International Journal of Human-Computer Studies*, *165*, 102835. doi:

Lin, C. J., & Cheng, C. F. (2022, 7). Modeling the effect of target shape on movement performance in a 1d2d fitts task. *Mathematics 2022, Vol. 10, Page 2568*, *10*, 2568. Retrieved from `https://www.mdpi.com/2227-7390/10/15/2568/htm` `https://www.mdpi.com/2227-7390/10/15/2568` doi:

MacKenzie, I. S. (1989, 9). A note on the information-theoretic basis for fitts' law. *Journal of Motor Behavior*, *21*, 323-330. doi:

Mackenzie, I. S. (2015). Fitts' throughput and the remarkable case of touch-based target selection. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9170*, 238-249. Retrieved from `https://link.springer.com/chapter/10.1007/978-3-319-20916-6_23` doi:

MacKenzie, I. S., & Buxton, W. (1992). Extending fitts' law to two-dimensional tasks. *Conference on Human Factors in Computing Systems - Proceedings*, 219-226. Retrieved from `https://dl.acm.org/doi/10.1145/142750.142794` doi:

MacKenzie, I. S., & Isokoski, P. (2008). Fitts' throughput and the speed-accuracy trade-off. *Conference on Human Factors in Computing Systems - Proceedings*, 1633-1636. Retrieved from `https://dl.acm.org/doi/10.1145/1357054.1357308` doi:

MacKenzie, I. S., Kauppinen, T., & Silfverberg, M. (2001). Accuracy measures for evaluating computer pointing devices. *Conference on Human Factors in Computing Systems - Proceedings*, 9-16. Retrieved from `https://dl.acm.org/doi/10.1145/365024.365028` doi:

Mackenzie, I. S., & Teather, R. J. (2012). Fittstilt: The application of fitts' law to tilt-based interaction. *NordiCHI 2012: Making Sense Through Design - Proceedings of the 7th Nordic Conference on Human-Computer Interaction*, 568-577. Retrieved from `https://dl.acm.org/doi/10.1145/2399016.2399103` doi:

Majaranta, P., Räihä, K.-J., Hyrskykari, A., & Špakov, O. (2019). Eye movements and human-computer interaction. *Eye movement research: An introduction to its scientific foundations and applications*, 971–1015.

Matsuyama, S., & Karashima, M. (2017). A study on characteristics of hand gesture pointing operation versus mouse pointing operation: A comparison of velocity waves of operation time between mouse pointing and hand gesture pointing with two kinds of control-display ratio. *Communications in Computer and Information Science*, *713*, 170-176. Retrieved from `https://link.springer.com/chapter/10.1007/978-3-319-58750-9_24` doi:

McAnally, K., & Wallis, G. (2024). Effects of auditory feedback on visually-guided movement in real and virtual space. *International Journal of Human–Computer Interaction*, *40*, 4431-4440. Retrieved from `https://www.tandfonline.com/doi/abs/10.1080/10447318.2023.2216091` doi:

McArthur, V., Castellucci, S. J., & MacKenzie, I. S. (2009). An empirical comparison of "wiimote" gun attachments for pointing tasks. *EICS'09 - Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, 203-208. Retrieved from `https://dl.acm.org/doi/10.1145/1570433.1570471` doi:

Natapov, D., & MacKenzie, I. S. (2010). The trackball controller: Improving the analog stick. *Future Play 2010: Research, Play, Share - International Academic Conference on the Future of Game Design and Technology*, 175-182. Retrieved from `https://dl.acm.org/doi/10.1145/1920778.1920803` doi:

Negi, S., & Mitra, R. (2020). Fixation duration and the learning process: an eye tracking study with subtitled videos. *Journal of Eye Movement Research*, *13*, 1-15. Retrieved from `/pmc/articles/PMC8012014/` `/pmc/articles/PMC8012014/?report=abstract` `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8012014/` doi:

Oehl, M., Sutter, C., & Ziefle, M. (2007). Considerations on efficient touch interfaces–how display size influences the performance in an applied pointing task. In *Symposium on human interface and the management of information* (pp. 136–143).

Olafsdottir, H. H., Guiard, Y., Rioul, O., & Perrault, S. T. (2012). A new test of throughput invariance in fitts' law: Role of the intercept and of jensen's inequality. *Interfaces, the quaterly magazine of BCS interaction group*(93), 8–pages.

Ortega, M., & Nigay, L. (2009). Airmouse: Finger gesture for 2d and 3d interaction. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *5727 LNCS*, 214-227. Retrieved from `https://link.springer.com/chapter/10.1007/978-3-642-03658-3_28` doi:

Pandey, L., & Arif, A. S. (2022, 11). Design and evaluation of a silent speech-based selection method for eye-gaze pointing. *Proceedings of the ACM on Human-Computer Interaction*, *6*, 328-353. Retrieved from `https://dl.acm.org/doi/10.1145/3567723` doi:

Park, K. S., Hong, G. B., & Lee, S. (2012, 5). Fatigue problems in remote pointing and the use of an upper-arm support. *International Journal of Industrial Ergonomics*, *42*, 293-303. doi:

Pino, A., Tzemis, E., Ioannou, N., & Kouroupetroglou, G. (2013). Using kinect for 2d and 3d pointing tasks: Performance evaluation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence* *and Lecture Notes in Bioinformatics)*, *8007 LNCS*, 358-367. Retrieved from https://link.springer.com/chapter/10.1007/978-3-642-39330-3_38 doi:

Rajanna, V., & Hammond, T. (2018, 6). A fitts' law evaluation of gaze input on large displays compared to touch and mouse inputs. *Proceedings - COGAIN 2018: Communication by Gaze Interaction*. Retrieved from https://dl.acm.org/doi/10.1145/3206343.3206348 doi:

Ramcharitar, A., & Teather, R. J. (2017, 5). A fitts' law evaluation of video game controllers: Thumbstick, touchpad and gyrosensor. *Conference on Human Factors in Computing Systems - Proceedings*, *Part F127655*, 2860-2866. Retrieved from https://dl.acm.org/doi/10.1145/3027063.3053213 doi:

Roig-Maimó, M. F., MacKenzie, I. S., Manresa-Yee, C., & Varona, J. (2017, 9). Evaluating fitts' law performance with a non-iso task. *ACM International Conference Proceeding Series*, *Part F131194*. Retrieved from https://dl.acm.org/doi/10.1145/3123818.3123827 doi:

Roig-Maimó, M. F., MacKenzie, I. S., Manresa-Yee, C., & Varona, J. (2018, 4). Head-tracking interfaces on mobile devices: Evaluation using fitts' law and a new multi-directional corner task for small displays. *International Journal of Human-Computer Studies*, *112*, 1-15. doi:

Roig-Maimó, M. F., & Mas-Sansó, R. (2023, 9). Is it worth the hassle moving from one-directional to multi-directional on fitts' law evaluation. *ACM International Conference Proceeding Series*. Retrieved from https://dl.acm.org/doi/10.1145/3612783.3612787 doi:

Sheikh, I. H., & Hoffmann, E. R. (1994). Effect of target shape on movement time in a fitts task. *Ergonomics*, *37*, 1533-1547. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/00140139408964932 doi:

Shojaeizadeh, M., Djamasbi, S., & Trapp, A. C. (2016). Density of gaze points within a fixation and information processing behavior. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9737*, 465-471. Retrieved from https://link.springer.com/chapter/10.1007/978-3-319-40250-5_44 doi:

Skaramagkas, V., Giannakakis, G., Ktistakis, E., Manousos, D., Karatzanis, I., Tachos, N., ... Tsiknakis, M. (2023). Review of eye tracking metrics involved in emotional and cognitive processes. *IEEE Reviews in Biomedical Engineering*, *16*, 260-277. doi:

Soukoreff, R. W., & MacKenzie, I. S. (2004, 12). Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in hci. *International Journal of Human-Computer Studies*, *61*, 751-789. doi:

Sundar, S. S., Bellur, S., Oh, J., Xu, Q., & Jia, H. (2014). User experience of on-screen interaction techniques: An experimental investigation of clicking, sliding, zooming, hovering, dragging, and flipping. *Human–Computer Interaction*, *29*, 109-152. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/07370024.2013.789347 doi:

Te'eni, D. (1992, 1). Analysis and design of process feedback in information systems: old and new wine in new bottles. *Accounting, Management and Information Technologies*, *2*, 1-18. doi:

Velloso, E., Alexander, J., Bulling, A., & Gellersen, H. (2015). Interactions under the desk: A characterisation of foot movements for input in a seated position. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9296*, 384-401. Retrieved from https://link.springer.com/chapter/10.1007/978-3-319-22701-6_29 doi:

Wang, Y., Yu, C., Qin, Y., Li, D., & Shi, Y. (2013). Exploring the effect of display size on pointing performance. *ITS 2013 - Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces*, 389-392. Retrieved from https://dl.acm.org/doi/10.1145/2512349.2514911 doi:

Weiss, H., Tang, J., Williams, C., & Stirling, L. (2024, 4). Performance on a target acquisition task differs between augmented reality and touch screen displays. *Applied Ergonomics*,

*116*, 104185. doi:
Wobbrock, J. O., Jansen, A., & Shinohara, K. (2011). Modeling and predicting pointing errors in two dimensions. *Conference on Human Factors in Computing Systems - Proceedings*, 1653-1656. Retrieved from `https://dl.acm.org/doi/10.1145/1978942.1979183` doi:
Wobbrock, J. O., Shinohara, K., & Jansen, A. (2011). The effects of task dimensionality, endpoint deviation, throughput calculation, and experiment design on pointing measures and models. In *Proceedings of the sigchi conference on human factors in computing systems* (p. 1639–1648). New York, NY, USA: Association for Computing Machinery. Retrieved from `https://doi.org/10.1145/1978942.1979181` doi:
Wurth, S. M., & Hargrove, L. J. (2014). A real-time comparison between direct control, sequential pattern recognition control and simultaneous pattern recognition control using a fitts' law style assessment procedure. *Journal of neuroengineering and rehabilitation*, *11*(1), 91.
Yamauchi, T. (2013). Mouse trajectories and state anxiety: Feature selection with random forest. *Proceedings - 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII 2013*, 399-404. doi:
Yu, D., Liang, H. N., Lu, X., Fan, K., & Ens, B. (2019, 11). Modeling endpoint distribution of pointing selection tasks in virtual reality environments. *ACM Transactions on Graphics*, *38*, 13. Retrieved from `https://dl.acm.org/doi/10.1145/3355089.3356544` doi:
Yuan, H., Calic, J., Fernando, A., & Kondoz, A. (2013). Investigation and evaluation of pointing modalities for interactive stereoscopic 3d tv. *Proceedings - IEEE International Conference on Multimedia and Expo*. doi:
Zhai, S., Kong, J., & Ren, X. (2004, 12). Speed–accuracy tradeoff in fitts' law tasks—on the equivalency of actual and nominal pointing precision. *International Journal of Human-Computer Studies*, *61*, 823-856. doi:
Zhai, S., Smith, B. A., & Selker, T. (1997). Improving browsing performance: A study of four input devices for scrolling and pointing tasks. *Human-Computer Interaction INTERACT '97*, 286-293. Retrieved from `https://link.springer.com/chapter/10.1007/978-0-387-35175-9_48` doi:
Zhang, X., & MacKenzie, I. S. (2007). Evaluating eye tracking with iso 9241 - part 9. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *4552 LNCS*, 779-788. Retrieved from `https://link.springer.com/chapter/10.1007/978-3-540-73110-8_85` doi:
Zhang, X., Zha, H., & Feng, W. (2012). Extending fitts' law to account for the effects of movement direction on 2d pointing. *Conference on Human Factors in Computing Systems - Proceedings*, 3185-3194. Retrieved from `https://dl.acm.org/doi/10.1145/2207676.2208737` doi:

## Appendix

## Appendix A. Description of Measures

### *A.1. Description of Cursor Movement Measures*

- **Target re-entry**: The number of occurrence when the pointer enters the target region, leaves, then re-enters the target region (unit: times)
- **Task axis crossing**: The number of occurrence when the pointer crosses the task axis on the way to the target (unit: times)
- **Movement direction change**: The number of occurrence when the pointer's path relative to the task axis changes direction (unit: times)
- **Orthogonal direction change**: The number of occurrence when the pointer's

path orthogonal to the task axis changes direction (unit: times)
- **Movement variability**: A continuous measure computed from the x-y coordinates of the pointer during a movement task that represents the extent to which the sample points lie in a straight line along an axis parallel to the task axis (unit: pixel)
- **Movement error**: The average deviation of the sample points from the task axis, irrespective of whether the points are above or below the axis (unit: pixel)
- **Movement offset**: The mean deviation of sample points from the task axis (unit: pixel)

### *A.2. Description of Eye Movement Measures*

It is important to note that some eye movement measures were calculated on a per-target basis.

- **Fixation duration**: The length of time the eyes remain relatively still per target (Total fixation time / Number of targets; unit: second)
- **Saccade duration**: The time taken for rapid eye movements between fixations per target (Total saccade time / Number of targets; unit: second)
- **Number of fixations**: The count of times the eyes pause to focus on specific points per target (Total number of fixations / Number of targets; unit: times)
- **Number of saccades**: The quantity of rapid eye movements between fixations per target (Total number of saccades / Number of targets; unit: times)
- **Saccade amplitude**: The distance traveled during a saccade (Total saccade amplitude / Total number of saccades; unit: pixel)
- **Saccade velocity**: The speed at which the eyes move during saccades (Total saccade amplitude / Total saccade time; unit: pixel/second)
- **Pupil diameter**: The size of the pupil (unit: milimeter)

## Author biographies

- **Yuhwa Hong** earned her B.S. in Industrial Engineering Design from KOREATECH and her M.S. in Industrial Engineering from Yonsei University, South Korea. She is currently a Ph.D. candidate in Industrial and Systems Engineering at KAIST in Daejeon. Her research interests include HCI and XR.
- **Haejun Kim** earned his B.S. in Industrial Engineering and B.A. in Philosophy from Hanyang University, Seoul, South Korea. He is currently pursuing an M.S. in Data Science at KAIST in Daejeon. His research interests include HCI, AI, and VR.
- **Jihae Yu** earned her B.S. in Electronics Engineering from Kangwon National University, Chuncheon, South Korea. She is currently pursuing an M.S. in Data Science at Kangwon National University. Her research interests include HCI and XR.
- **Heedo Shin** is currently pursuing a Bachelor's degree in Computer Science and Engineering at Kangwon National University in Chuncheon, South Korea. His research interests include HCI and XR.
- **Xiaoqun Yu** earned his Ph.D. in Industrial and Systems Engineering from KAIST in Daejeon, South Korea. He is currently an Assistant Professor in the Department of Mechanical and Industrial Design at Southeast University in Nan-

jing, China. His research interests include HCI, AI, and healthy aging.

- **Shuping Xiong** earned his Ph.D. in Industrial Engineering from HKUST, Hong Kong, China. He is currently a Professor in Industrial and Systems Engineering at KAIST, South Korea, and previously served as a Visiting Professor at Purdue University, United States. His research focuses on HCI, ergonomics, healthy aging, and human performance modeling.
- **Woojoo Kim** earned his Ph.D. in Industrial and Systems Engineering from KAIST in Daejeon, South Korea. He is currently an Assistant Professor in Liberal Studies at Kangwon National University. His research interests include HCI, XR, and 3D UIs.