Estimating the Effects of Encumbrance and Walking on Mixed Reality Interaction

Tinghui Li
School of Computer Science
The University of Sydney
Sydney, New South Wales, Australia
tinghui.li@sydney.edu.au

Eduardo Velloso
School of Computer Science
The University of Sydney
Sydney, New South Wales, Australia
eduardo.velloso@sydney.edu.au

Anusha Withana
School of Computer Science
The University of Sydney
Sydney, New South Wales, Australia
anusha.withana@sydney.edu.au

Zhanna Sarsenbayeva
School of Computer Science
The University of Sydney
Sydney, New South Wales, Australia
zhanna.sarsenbayeva@sydney.edu.au

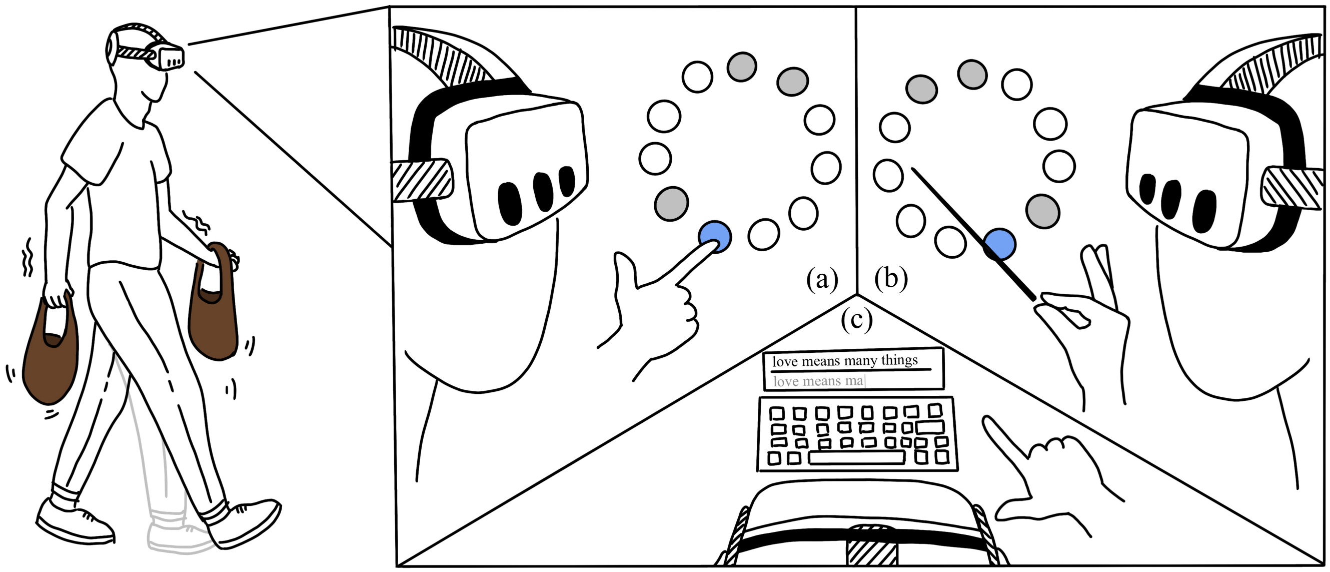


Figure 1: User is interacting with the mixed reality objects while carrying bags and walking. The tasks on the right are: (a) target acquisition (direct selection); (b) target acquisition (ray-cast); (c) text entry.

Abstract

This paper investigates the effects of two situational impairments—encumbrance (i.e., carrying a heavy object) and walking—on interaction performance in canonical mixed reality tasks. We built Bayesian regression models of movement time, pointing offset, error rate, and throughput for target acquisition task, and throughput, UER, and CER for text entry task to estimate these effects. Our results indicate that 1.0 kg encumbrance increases selection movement time by 28%, decreases text entry throughput by 17%, and increase UER by 50%, but does not affect pointing offset. Walking led to a 63% increase in ray-cast movement time and a 51% reduction in text entry throughput. It also increased selection pointing offset by 16%, ray-cast pointing offset by 17%, and error rate by 8.4%. The interaction effect on 1.0 kg encumbrance and walking resulted in a 112% increase in ray-cast movement time. Our findings enhance the understanding of the effects of encumbrance and walking on mixed reality interaction, and contribute towards accumulating knowledge of situational impairments research in mixed reality.

CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713492

CCS Concepts

• **Human-centered computing → Mixed / augmented reality**; **Empirical studies in HCI**; **Empirical studies in accessibility**.

Keywords

Situational Impairments, Encumbrance, Walking, Fitts's Law, Text Entry

ACM Reference Format:
Tinghui Li, Eduardo Velloso, Anusha Withana, and Zhanna Sarsenbayeva. 2025. Estimating the Effects of Encumbrance and Walking on Mixed Reality Interaction. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 24 pages. https://doi.org/10.1145/3706598.3713492

1 Introduction

The recent broader commercialisation of Mixed Reality (MR) headsets (e.g., Meta Quest 3, Apple Vision Pro) has led to this technology finally getting out of the confines of research labs and out into the real world—with all of its complexity. Early adopters have been spotted using these technologies in settings as diverse as the gym[1], aeroplanes[2], and even driving[3]. Much research on MR interaction techniques has been conducted in "ideal" circumstances, removing extraneous sources of variability that might affect results—participants sit or stand still, without the burden of carrying objects, and the task has their full attention. Though these experiments are critical for understanding the upper bounds of performance, they must be complemented by studies subject to less-than-ideal conditions. This is particularly significant because prior research has shown that contextual and environmental factors—known as *situational impairments* [67]—negatively affect interaction with technology, including desktop computers [55], smartphones [58, 59, 62] and wearable displays [22, 93].

To seamlessly integrate MR technologies into everyday tasks, it is crucial to understand how situational impairments can affect their performance. In this paper, we investigate the effects of two such situational impairments—**encumbrance** (i.e., carrying a weighted object) and **walking**—on the performance of canonical MR tasks. Many users interact with digital devices while on the move; thus, if future MR usage resembles current mobile phone use, we can expect that people will use them while walking. Similarly, whether commuting, shopping, or multitasking, users often find themselves using digital devices while holding or carrying bags and other objects. Hence, understanding the effects of encumbrance and walking is essential for designing MR interfaces that accommodate real-world scenarios and enhance usability for individuals in diverse, physically constrained situations.

In this study, we simulated encumbrance using wrist-attached bracelets with adjustable weights (three levels: no weight, 0.5 kg, and 1.0 kg) and constant walking speed to simulate the walking condition similar to previous studies [46]. We used 2D serial pointing tasks commonly used in Fitts' Law experiments (i.e., target acquisition task via direct selection and target acquisition task via ray-casting) and text entry task (Figure 1). We measured the effects of encumbrance and walking using movement time [46, 47], pointing offset [46, 59, 60, 62], error rate [25], and throughput [39] in the target acquisition tasks, and throughput [91], uncorrected error rate (UER) [70, 91], and corrected error rate (CER) [70, 91] in the text entry task (we also report words-per-minute in the Appendix A). Our results indicate that being encumbered with 1.0 kg increased selection movement time by 28%, decreased selection throughput by 22%, decreased text entry throughput by 17%, and increase uncorrected error rate by 50%, but did not affect pointing offset and error rate. Walking led to a 63% increase in ray-casting movement time, a 32% decrease in ray-casting throughput, and a 51% reduction in text entry throughput. Pointing offset increased by 17%, and the error rate increased by 8.4% using the ray-cast interaction method while walking. The corrected error rate increased by 68%. The combined effect of 1.0 kg weight and walking increased movement time by 112% in the ray-casting target acquisition task and decreased text entry throughput by 58%.

Overall, our study contributes to the growing research area on situational impairments in MR by considering the encumbrance and walking situations when interacting in an MR context. The contributions of this paper are as follows:

- We build Bayesian regression models to quantify by how much encumbrance and walking increase movement time and decrease throughput. We show that walking has a greater impact than encumbrance and that pointing offset and error rate remain unaffected when encumbered, whereas walking affects both;
- We enhance the understanding of the effects of encumbrance and walking on MR interaction and contribute towards accumulating knowledge in situational impairments research expanding it to MR;
- We highlight the importance of considering situational impairments in MR interaction and propose potential strategies and directions to mitigate the impacts of situational impairments.

2 Related Work

Previous research has shown how different environmental and contextual factors influence the way we interact with technology. For example, cold environments [21, 58, 88], background noise [62], the user's mobile state [20], stress [59], and physical encumbrance [45, 46]—can adversely influence mobile interaction, interaction with desktop computers [55], and smartwatches [22]. These phenomena are known in the literature as Situationally-Induced Impairments and Disabilities (SIIDs), or *situational impairments* for short [67]. SIIDs are distinct from health-related impairments and disabilities because they are caused by external environmental factors, not the individual's physical or mental health [82]. Once the external factors change or are removed, the situational impairments also disappear, restoring the individual's capabilities. As we use technology under various contextual factors, situational impairments are bound to affect our interaction with these devices [82]. While we already knew the effects of situational impairments on mobile devices [61, 74], desktop computers [55], and wearable devices [10, 22, 89], their effects on interaction with MR headsets remain under-explored. Hence, in our study, we contribute to the growing body of research on situational impairments by quantifying the effect of encumbrance and walking on interaction with MR headsets and measuring changes to user performance while completing canonical MR tasks.

[1] https://www.essentiallysports.com/bodybuilding-news-bodybuilder-wears-apple-vision-pro-while-working-out-in-the-gym-and-heres-what-happened/ [Accessed: 2024-09-11]
[2] https://www.tomsguide.com/computing/vr-ar/i-flew-8000-miles-wearing-apple-vision-pro-heres-what-its-really-like [Accessed: 2024-09-11]
[3] https://gizmodo.com.au/2024/02/here-are-the-wildest-ways-people-have-been-using-their-apple-vision-pro/ [Accessed: 2024-09-11]

2.1 Encumbrance and Walking as Situational Impairments

Encumbrance can occur in various everyday scenarios where individuals find themselves carrying different items and objects that prevent or limit the dexterity and mobility of their hands. Prior research mostly studied the effects of encumbrance on mobile interaction [46, 47]. Findings from prior research showed that encumbrance negatively affects one-handed interaction with smartphones, increasing the error rate by 40% [46]. The overall mean movement times increase by 18% and 12% for tapping and dragging interaction methods, respectively [47]. Though encumbrance can also hinder MR interaction and negatively affect user performance, these effects are yet to be quantified.

Further, prior research has mostly focused on the effects of walking while interacting with smartphones [46]. In a study by Lin et al. [33], participants completed standard target selection tasks while seated and while walking with obstacles. The study found that navigating an obstacle course posed a greater challenge for users compared to the seated condition when completing tapping tasks on mobile devices [33]. Similarly, Schildbach and Rukzio [64] found that walking negatively impacts performance in target selection tasks and reading comprehension. The authors further proposed that the detrimental effects of walking on target selection could be mitigated by increasing target sizes [64]. Hence, it is likely that walking also negatively affects user performance when completing tasks in MR, but no previous studies have quantified this effect.

2.2 Measuring Performance in MR

2.2.1 Target Acquisition. Most MR applications use either of two target acquisition techniques [4]: direct selection and ray-casting. Direct selection involves choosing objects by directly touching them. In contrast, ray-casting allows users to select objects from a distance by projecting a visible "ray" (line) from the point of origin (e.g., the user's hand). Direct selection has better throughput [6] and avoids issues of precision and parallax associated with angular control [52]. However, the selection area is limited and requires more pronounced gestures or hand movements from the user [65]. In contrast, ray-casting allows one to perform gestures on a larger selection area and manipulate objects located at a further distance without having to perform pronounced movements as compared to direct touch [4]. However, it has also been shown to be susceptible to tremor and parallax effects [28, 52]. Therefore, we used two target selection tasks: target acquisition via direct selection and target acquisition via ray-casting. Fitts' law is commonly used to model the performance during target acquisition tasks [37]. The task layout (i.e., Fitts' ring, see Figure 3a and 3b) is used in Human-Computer Interaction (HCI) research to evaluate user performance [59, 62] and propose new models [85]. It contains a sequence of targets located on the circumference of a ring, and users are required to reach the targets one at a time. Common performance measures used in pointing tasks include movement time, pointing offset, error rate, and throughput, which we adopt in our study.

Movement Time: Movement time is the duration it takes for the user to select a specific target [16, 47, 85]. It quantifies the time required to move to a target area as a function of the distance to the target and the size of the target [16]. In MR scenarios, this might involve the time taken to navigate through menus, select items, or execute commands. In our study, we computed movement time as the time between target selections.

Pointing Offset: This corresponds to the distance between the end-point location (i.e., where the user selects or reaches) and the centre of the target [46, 59, 62]. As such, it is an error measure, quantifying how close the user's final action is to the intended target's centre. Situational impairments tend to increase this pointing offset [46–48].

Error Rate: Different from pointing offset, error rate measures the frequency of mistakes made during target selection [59, 62]. This includes uncorrected selections, wrong selections, failed attempts to engage with virtual elements or misinterpretations of interface prompts. A lower error rate suggests that the user can effectively interact with the MR system [46, 47].

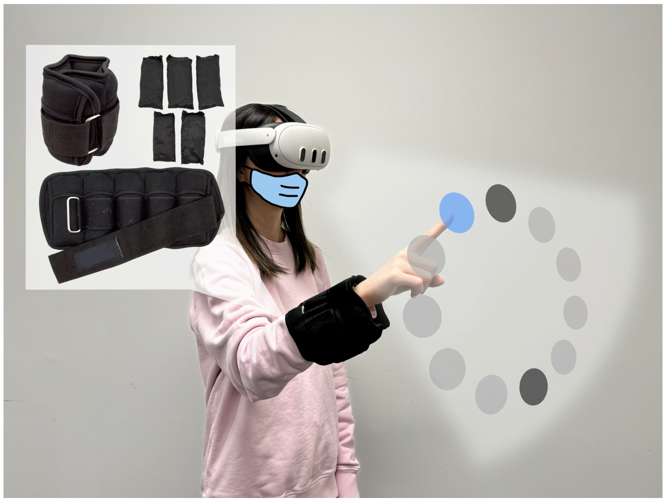
Throughput: Throughput, in bits per second (bits/s), combines speed and accuracy in a single measure computed over repeated trials [39]. Speed is represented by the time to complete a task, usually known as movement time in Fitts' law experiments. Throughput also considers accuracy, often measured by the deviation from the target.

2.2.2 Text Entry. Text entry is another commonly used task in MR. Previous research has shown that the task can be successfully used to analyse user performance [59, 62]. The task usually presents a set of phrases to be typed verbatim with or without the time limit [43]. Every user has their own internal subjective speed-accuracy bias, which may change with purpose and context [91]. Thus, separate measures of speed and accuracy will vary under different conditions [91]. In light of this, we used a robust performance measure, *throughput*, proposed by Zhang et al. [91], that effectively combines the information provided by speed and accuracy.

Throughput: Throughput combines both speed and accuracy to provide a comprehensive measure of interaction performance in text entry tasks [12, 13, 79, 91]. It is calculated by considering the number of successfully completed tasks within a given time frame while accounting for any errors made. Zhang et al. [91] proposed an independent throughput metric based on Shannon information theory [68] that takes into account both uncorrected error rate (UER) and corrected error rate (CER), and character per second entry rate [70, 91]. We adapt this new throughput metric in this paper to evaluate the text entry task. High throughput reflects a balance of quick and accurate interactions, demonstrating that the user can maintain effective performance [12, 79, 91].

UER and CER: Uncorrected errors are mistakes that remain in the transcribed text [70, 91], representing the wrongly typed characters that the user does not correct. Corrected errors, on the other hand, are those that are made but fixed during the entry process (e.g., using backspace) [70, 91], representing wrongly typed characters that were fixed during the text entry task.

To summarise, based on the literature, we use three tasks: target acquisition via direct selection, target acquisition via ray-cast, and text entry (Figure 3) in this work to quantify the effects of encumbrance and walking on user performance. Based on previous research [44, 59, 62], we operationalise performance with the four main measures from target acquisition task and three main measure from text entry task to quantify the effects of encumbrance and



(a) Participant is encumbered with wrist-attached bracelet and performing target acquisition (direct selection) task. We use wrist-attached bracelets to represent bags with corresponding weights.

(b) Participant is walking and performing target acquisition (direct selection) task. One researcher in the front represents the pacesetter to control the walking speed, and the participant follows the pacesetter's speed.

Figure 2: Example of the experimental conditions.

walking: movement time, pointing offset, error rate, target acquisition throughput, text entry throughput, UER, and CER. The WPM was also analysed in the Appendix A. These measures have also been used in previous work on the effects of SIIDs [44, 59, 62].

3 Method

In this study, we investigate the effect of encumbrance and walking on three common MR interaction tasks: target acquisition via direct selection, target acquisition via ray-casting, and text entry. We conducted a 3 × 2 within-subjects design: weight (0 kg, 0.5 kg, 1.0 kg) × movement (standing still, walking at a constant speed). We conducted the study in a controlled environment (5m × 6m room) and simulated encumbrance and walking conditions in order to exclude other factors that might potentially interfere with the experimental results. The study received ethics clearance from the Human Research Ethics Committee (HREC) of the University of Sydney (Application number: 2019/553).

3.1 Conditions

3.1.1 Encumbrance. We used wrist-worn bracelets with adjustable weights to operationalise encumbrance by attaching them to the participant's dominant hand (Figure 2a). Though encumbrance can be manifested in many ways, in this study, we tried to isolate the effect of weight. This prioritises the study's internal validity at the expense of other interesting effects, such as the momentum of swinging bags. We used the following weights: 0 kg (no bracelet), 0.5 kg (e.g., a standard bottle of water of 500 ml or a small handbag) and 1.0 kg (e.g., a textbook or a mid-sized laptop).

3.1.2 Walking. We asked our participants to walk along a predefined path in a clockwise direction, as shown in Figure 2b. We followed previous research [29, 46], where the walking speed was set to the participants' preferred daily walking speed. A participant first walked, setting their desired walking speed, and the researcher followed their speed. Once the researcher had grasped the speed, the participant followed the researcher to do the actual user study. We made this choice to maximize impairing walking effects, as users were not able to slow down if the task became difficult [29]. We chose not to control the walking speed of the participants (e.g., using a treadmill) to let participants find a balance between input performance and walking speed [3, 30, 36, 45]. In movement conditions, participants only stopped walking when they finished the task.

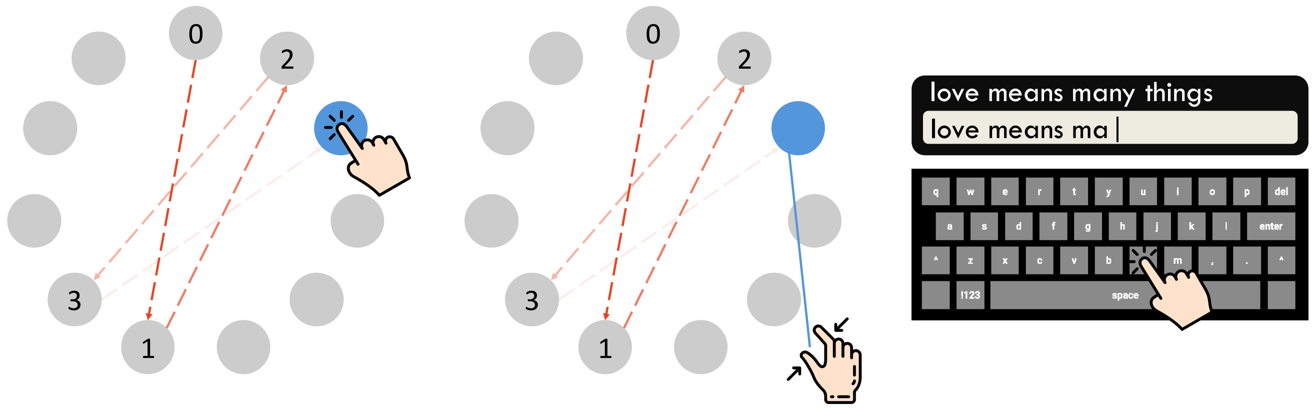
3.2 Tasks

We developed a custom MR application for the Meta Quest 3 in Unity `2022.3.12f1` and presented these tasks to participants in a counterbalanced order to avoid sequence effects. We describe each task in detail below.

3.2.1 Target Acquisition – Direct Selection. In this task, we asked users to select the target by tapping it with the index finger of their dominant hand. Following previous work [85, 87, 92], we based the task layout on the Fitts' ring task.

The task involves 11 targets arranged in a circle, with a starting position in the centre. First, participants need to select a start target located at the centre of the Fitts' ring circle. Then, participants selected the targets in an alternating sequence as they changed color (blue), indicating the next target (Figure 3a).

The start target and the movement towards the first selection ('0' in Figure 3a) were excluded from the analysis, as they do not involve the same movement amplitude or attack velocity as the subsequent selections [37]. In our study, distance and size of targets were covariates, not independent variables, allowing us to improve



(a) Target Acquisition (Direct Selection) (b) Target Acquisition (Ray-Casting) (c) Text Entry

Figure 3: The visual representation of the tasks used in the experiment: (a) target acquisition (direct selection); (b) target acquisition (ray-casting); (c) text entry. Blue circle represents the target that needs to be selected. Participants followed the path indicated by the red dashed arrows to select the targets sequentially.

the precision of our performance estimates (e.g., movement time, pointing offset) and draw comparisons between levels of encumbrance and walking. However, we are not interested in the causal effect of these variables as they are already well-understood.

We followed the Meta developer guidelines[4] and set the distance to the Fitts' ring to be 49 cm away from the central lens of the headset. The width/diameter and distance between the targets also followed the guidelines. We chose three target sizes: 4 cm, 6 cm, and 8 cm (diameters) with a corresponding distance of 3 cm and 5 cm between them. This resulted in a total of six combinations (i.e., three levels of diameters × two levels of distances) for the Fitts' task. These combinations of width and distance allowed us to vary the index of difficulty (ID) for the target acquisition (direct selection) task. Participants had to go through all 11 targets, repeating the process a total of 6 times for each of the IDs. The corresponding IDs and amplitudes are listed in Table 1.

Table 1: The corresponding ID and amplitude for different width/diameter and distance in target acquisition (direct selection) task.

D \ W	ID			Amplitude		
	4 cm	6 cm	8 cm	4 cm	6 cm	8 cm
3 cm	0.81	0.58	0.46	28.10°	35.64°	42.92°
5 cm	1.17	0.87	0.70	35.64°	42.92°	49.82°

3.2.2 Target Acquisition – Ray-Casting. The target acquisition via ray-casting task asked participants to select a target by projecting a visible "ray" (line) from a point of origin (user's hand) as shown in Figure 3b. We asked participants to pinch to select the targets using the casting ray in an alternating sequence. First, they had to pinch to select the start target located at the centre of the Fitts' ring circle. Once they did so, the targets changed colour (blue) in alternating sequences, and participants needed to reach them one by one (in Figure 3b). The task also used Fitts' ring [85, 92]. We set the distance to the Fitts' ring to be 200 cm from the central lens of the headset as recommended by the Meta developer guidelines. We further followed the guidelines to design targets of three sizes: 8 cm, 12 cm, and 16 cm diameters, with a corresponding distance of 6 cm and 10 cm between targets, resulting in a total of six combinations (i.e., three levels of diameters × two levels of distances) for the Fitts' task. Similar to the previous task, we created 6 IDs, and participants had to go through all 11 targets. The corresponding IDs and amplitudes are listed in Table 2.

Table 2: The corresponding ID and amplitude for different width/diameter and distance in target acquisition (ray-casting) task.

D \ W	ID			Amplitude		
	8 cm	12 cm	16 cm	8 cm	12 cm	16 cm
6 cm	0.81	0.58	0.46	13.98°	17.92°	21.80°
10 cm	1.17	0.87	0.70	17.92°	21.80°	25.64°

3.2.3 Measures of Target Acquisition. We followed prior research on SIIDs [59, 62] and used movement time, pointing offset, error rate, and throughput to measure performance during target acquisition tasks.

Movement time (MT) is recorded as the time (in milliseconds) taken by participants to select the target. MT is linearly associated with the index of difficulty (ID) [38], as shown in Equation 1, where ID depends on the target width (W) and distance (D) [38] as shown in Equation 2, a and b are empirically determined coefficients. The smaller the movement time, the better the performance.

$$MT = a + b \cdot ID \tag{1}$$

[4] https://developer.oculus.com/resources/hands-design-ui/ [Accessed: 2024-09-11]

$$ID = \log_2\left(\frac{D}{W} + 1\right) \tag{2}$$

We used the pointing offset to evaluate the precision of the end-point location in target acquisition tasks. We measured the pointing offset of the end-point location as the distance between the target centre (x_0, y_0, z_0) and the actual location of user selection $(\hat{x}, \hat{y}, \hat{z})$ as shown in Equation 3. The larger the pointing offset, the worse the performance.

$$\text{Offset} = \sqrt{(\hat{x} - x_0)^2 + (\hat{y} - y_0)^2 + (\hat{z} - z_0)^2} \tag{3}$$

We measured the error rate [90] as the percentage of the wrongly selected targets to the total number of targets (Equation 4). The smaller the error rate, the better the performance.

$$E_{target} = \frac{\text{Incorrect Selections}}{\text{Total Targets}} \cdot 100\% \tag{4}$$

We further evaluated throughput (TP) for target acquisition tasks, which combines speed and accuracy in a single measure computed over repeated trials [39]. The throughput is shown in Equation 5, where ID_e represents the effective index of difficulty. ID_e is calculated using Equation 6, where W_e is the effective target width [83]. The measure W_e is calculated from the standard deviation of the selection coordinates collected across a series of trials, which is $W_e = 4.133 \times SD$. The smaller the throughput, the worse the performance.

$$TP = \frac{ID_e}{MT} \tag{5}$$

$$ID_e = \log_2\left(\frac{D}{W_e} + 1\right) \tag{6}$$

3.2.4 Text Entry. In the text entry task, the application presented a sequence of phrases, one at a time, and we asked participants to type them verbatim in the text box. We followed previous studies [14, 72], in which 5 sentences were randomly selected from MacKenzie's phrase set [40] for participants to type under each of the conditions. The texts were of the same difficulty, ensuring fairness across conditions and removing any potential confounding effect of the presented text. We allowed participants to submit their sentences with wrong characters as restricting responses to fully correct sentences would lead to a loss of data on uncorrected error rate [70, 91]. Since we measure the throughput, both UER and CER were accounted for evaluation.

3.2.5 Measures of Text Entry. We used throughput to measure participants' performance in the text entry task. Text entry throughput is a robust performance measure for text entry that conveys the information found in speed and accuracy measures while also remaining stable across various speed-accuracy biases [91]. Characters per second (CPS) is the speed metric. The calculation is shown in Equation 7, where $I(X,Y)$ is the transmitted information that can be calculated as $I(X,Y) = H(X) - H_Y(X)$, with $H(X)$ representing the source information, and $H_Y(X)$ representing the conditional entropy. We report the WPM in Appendix A as supporting data.

$$\text{Throughput} = I(X,Y) \cdot CPS \tag{7}$$

We also used uncorrected error rate (UER) and corrected error rate (CER) to measure the performance in the text entry task. Uncorrected errors refer to mistyped characters that the user fails to correct during the text entry process (Equation 8). In contrast, corrected errors denote mistyped characters that the user successfully rectifies during the task (Equation 9).

$$UER = \frac{\text{Uncorrected and Not Fixed Character}}{\text{Total Character}} \cdot 100\% \tag{8}$$

$$CER = \frac{\text{Uncorrected but Fixed Character}}{\text{Total Character} + \text{Uncorrected but Fixed Character}} \cdot 100\% \tag{9}$$

3.3 Participants

We recruited 30 participants (16 men, 14 women) in our study (2 left-handed) using our university's notice board and snowball recruitment. Participants' average age was 24 (min = 19, max = 29, SD = 3). Prior to analysing the data of these participants, we followed a common practice to mirror the data for left-handed participants [62]. This is to ensure the axis and unit of the data were the same across all participants. 13 out of 30 participants were not familiar with the headset. The average height of our participants is 170 cm (min = 158 cm, max = 192 cm, SD = 9 cm); the average weight is 65 kg (min = 45 kg, max = 89 kg, SD = 14 kg). Our participants had exercised an average of 8 days (min = 0 day, max = 26 days, SD = 7 days) over the past month. Our sample covers a wide range of participants, from non-exercisers to daily exercisers. Each participant was given a unique anonymous ID (Participant ID) in our study. Our sample size is along the sample size standards for HCI research [8].

3.4 Procedure

Our study followed a 3 × 2 within-subjects design: weight (0 kg, 0.5 kg, 1.0 kg) × movement (standing still, walking at a constant speed). The order of conditions was randomly presented to the participant. This way, we minimised the impact of any potential fatigue or learning effects. Upon their arrival at our lab, we provided an overview of the study's purpose. Once participants understood the study and agreed to take part, they signed the consent form. We then asked participants to complete a background questionnaire. The background questionnaire collected participants' demographic information on gender, age, dominant hand, height, weight, frequency of exercise, and experience with MR. We collected this information to ensure our participants represented a broad range, and we used a within-subject design to reduce variability due to individual differences and systematic biases.

We then trained our participants using a tutorial with the three tasks until they were comfortable and familiar with the headset and each task to minimise learning effects. The tutorial tasks were shorter versions of the actual tasks: the target acquisition (direct selection) and target acquisition (ray-casting) tasks included only one trial of the Fitts' ring, and the text entry task involved a single sentence not included in the set of phrases used in the study.

After the training, we started the video recording and ended it after the interview finished. Participants completed the three tasks following the researcher's instructions. Participants were given time to rest between each condition to avoid fatigue and dizziness; this time was between 1-2 minutes and was determined by the

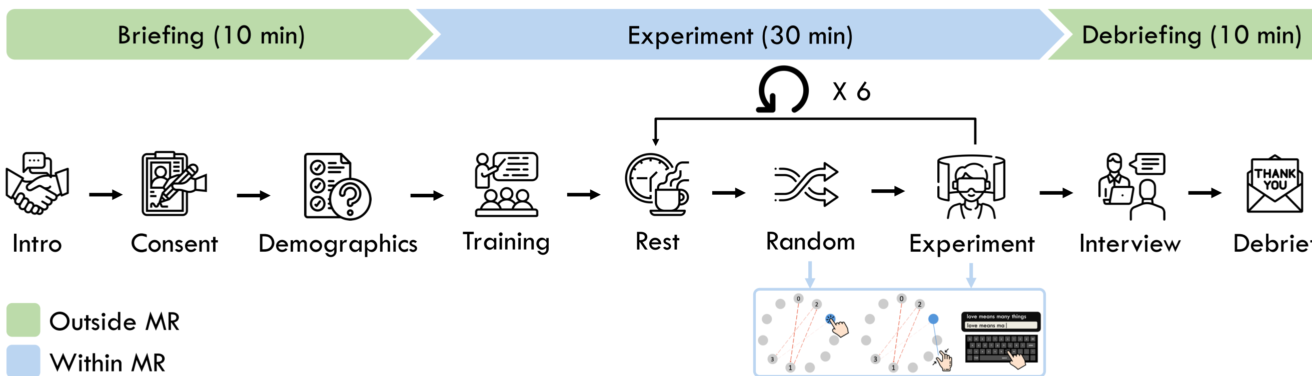


Figure 4: The experimental procedure.

participant's perception to ensure they felt rested enough. At the end of the experiment, we conducted a semi-structured interview with each participant to gain more insights into their experience during the study. The total duration of the study was 50-60 minutes per participant. The complete process is listed in Figure 4.

3.5 Data Analysis

We summarise our theoretical claims as a direct acyclical graph in Figure 5. We argue that encumbrance and walking affect movement time, pointing offset, error rate, target acquisition throughput, text entry throughput, UER, and CER at different levels. We further evaluate the interaction effect of encumbrance and walking. We also add covariates (i.e., ID, ID_e, and Width) known to affect the movement time, pointing offset, error rate, and throughput to increase the precision of our estimates. The order of the condition (Order) and the exercise frequency (Exercise) are kept in the models as fixed effects to control for fatigue and learning effects across individuals. Since we used a within-subject design with a random order of conditions, these do not confound our results [1, 19]. However, in scenarios with cumulative effects (e.g., fatigue), modelling Order as a fixed effect better estimates the trend without assuming random variation [53, 69]. We modelled participant ID as random effects reflecting the hierarchical structure of the data. The variables in the model are listed below:

- **Weight:** A numeric variable indicating the weight attached to the wrist in kg.
- **Motion:** A binary variable indicating whether the participant is walking or not.
- **ID:** A numeric variable indicating the index of difficulty. This covariate was used to analyse movement time and error rate [16, 46, 47].
- **$\mathbf{ID_e}$:** A numeric variable indicating the effective index of difficulty. This covariate was used to analyse target acquisition throughput [39].
- **Width:** A numeric variable indicating the width/diameter of the target. This covariate was used to analyse pointing offset [38].
- **Order:** A numeric variable indicating the order of the condition, which was treated as a fixed effect.
- **Exercise:** A numeric variable indicating the number of days to describe participants' exercising frequency within one month, which was treated as a fixed effect.
- **Participant ID:** A random effect used to model individual differences.

We employed Bayesian statistical methods in our analysis due to their enhanced flexibility, capacity to quantify uncertainty, and ability to facilitate future work to build upon it [5, 42]. This method is widely used in HCI research [5, 66]. We fit our models using the `brms` package [7], which implements Bayesian multilevel models in R using the Stan probabilistic programming language [9]. We used regularising priors designed to be sceptical of implausibly large effect sizes. We assessed the convergence and stability of the Markov Chain Monte Carlo sampling with R-hat, which should be lower than 1.01 [76] and the Effective Sample Size (ESS), which should be greater than 1000 [7]. All of our estimates fit these criteria. We report the posterior means of parameter estimates, the error of these estimates, and the upper and lower bounds of the 95% compatibility interval (i.e., credible interval, CI) [7]. This compatibility interval indicates the range of values where there is a 95% probability that the true value falls within. For full transparency, all our analysis scripts and results can be found in the supplementary material.

We report the hypothesis test results using Bayes Factor, which compares the likelihood of the observed data under the proposed model over the null condition. We interpret these values following the approach by Russo [57] and Wagenmakers et al. [78], considering values above 1 as supporting a given hypothesis, values under 3 offering anecdotal evidence; under 10, substantial evidence; under 30, strong evidence; under 100, very strong evidence; and above 100, extreme evidence. We note that p-values are not used in Bayesian statistics, and no claims about "statistical significance" should be derived from our results.

Additionally, we report the hypothesis test results using posterior probability as well, which reflects the updated belief about a hypothesis of the positive effect after seeing the data, considering both the prior and the likelihood [31]. We interpret the values following the approach by Wadinambiarachchi et al. [77], considering values higher than 90% as accepting the hypothesis; 60% - 90% as indicating some evidence, but it may not be strong enough;

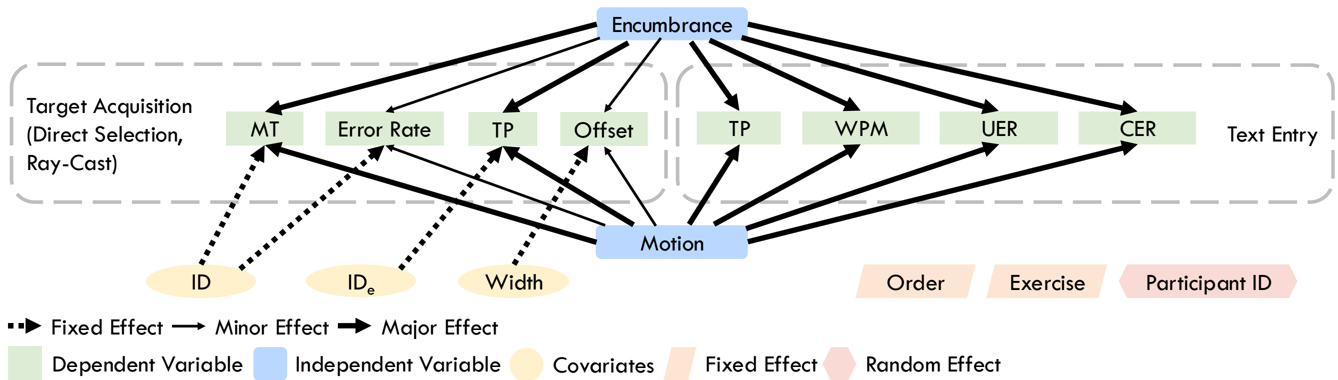


Figure 5: Theorised causal directed acyclic graph. The thickness of the arrow represents the effect of the factor. The thicker the arrow, the stronger its effect.

40% - 60% as suggesting the hypothesis is unlikely; 10% - 40% as indicating some evidence about negative effect, but it may not be strong enough; and below 10% as indicating strong evidence about negative effect.

We use Python `3.12.0` and R `4.3.1` to clean and analyse the data. We used `Pandas 2.1.2`, `Numpy 1.26.1` (Python), and `brms 2.21.0`, `ggplot2 3.4.4` (R) libraries for our data analysis. The qualitative analysis presented in Section 4.1.5 and Section 4.2.3 reflects on the participants' comments and behaviour during the study and their answers in the semi-structured interviews. We semantically analysed the video recordings from each participant and summarised their interview answers.

4 Results

4.1 Target Acquisition

We collected a total of 26,028 records in the target acquisition task. We filtered out unintentional selections by removing the targets that took less than 100 milliseconds to select from the dataset as accidental activations [46], leaving us with a total number of 24,886 records (11 targets/task/person × 6 conditions (3 weight × 2 motion) × 6 IDs × 2 tasks (direct selection/ray-casting) × 30 participants + wrong selections).

4.1.1 Movement Time. To analyse movement time, we filtered our data and kept only the data with correctly selected targets; uncorrected selections were used to analyse the error rate. The movement time follows a shifted Log-normal distribution [49, 54]. We used Bayesian regression to model the effect of encumbrance and walking on movement time in a log scale.

Table 3 reports the posterior means of parameter estimates, the errors of these estimates, and the upper and lower bounds of the 95% compatibility/credible interval (CI). The model suggests that both the effect of weight and walking have a 100% probability of leading to higher movement time for both target acquisition tasks. The Bayes Factor of 7.9×10^{308} for direct selection and 1.4×10^{38} for ray-casting suggest extreme support for weight leading to higher movement time, and the Bayes Factor of 4.9×10^{272} for direct selection and ∞ for ray-casting suggest extreme support for walking leading to higher movement time. The coefficients for weight (0.41, SD = 0.05, CI = [0.31, 0.51]) and motion (0.17, SD = 0.05, CI = [0.08, 0.27]) are positive for direct selection, which indicates an increase in movement time. In addition, the coefficients for weight (0.27, SD = 0.07, CI = [0.14, 0.41]) and motion (0.09, SD = 0.06, CI = [-0.03, 0.22]) are positive for ray-casting as well, which indicates an increase in movement time. We found no interaction effect of weight and motion on direct selection (mean = -0.07, error = 0.07, CI = [-0.21, 0.07]), where the probability is 19%. However, there is a substantial interaction effect of weight and motion on ray-casting (0.22, SD=0.10, CI = [0.02, 0.41]), where the probability is 93%.

Figure 6 illustrates the posterior means of parameter estimates and the corresponding bounds of the 95% CI for weight and motion. We can see that all factors increase the movement time, with ray-cast under walking condition having the highest effect on movement time. In summary, our model suggests that, on average, heavy weight and walking add towards increase in movement time. The effects of weight and walking in movement time on direct selection is less than ray-casting. The effect of weight in movement time is more than walking under direct selection, while the effect of walking in movement time is more than weight under ray-casting. Additionally, as the weight increases, its effect on movement time becomes more profound.

4.1.2 Pointing Offset. Similar to the movement time, we filtered our data and kept only the data with correctly selected targets; uncorrected selections were used to analyse the error rate. The pointing offset follows a shifted Log-normal distribution [49, 54]. We used Bayesian regression to model the effect of encumbrance and walking on pointing offset in log scale.

Table 4 reports the user's pointing offset for both direct selection and ray-casting tasks. It shows that the coefficients for weight (-0.25, SD = 0.07, CI = [-0.39, -0.11]) under direct selection indicate a decrease in the pointing offset, and the coefficients for weight (0.00, SD = 0.07, CI = [-0.14, 0.13]) under selection using ray-cast indicate a slight increase in the pointing offset. The effect of weight has a 0% probability of leading to larger pointing offset in direct selection and a 49% probability of leading to larger pointing offset in ray-cast. The Bayes Factor of 1.353 indicates no support for weight, leading to a smaller pointing offset in direct selection. Besides, the Bayes Factor

Table 3: Summary of the movement time model: MT ~ 1 + (1|participant) + order + exercise + ID · (weight · movement). We provide the posterior means of parameter estimates (Estimate), posterior error of these estimates (Error), and the upper and lower bound of the 95% CI. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00. The complete set of parameter estimates is in Table 10 in the Appendix B.

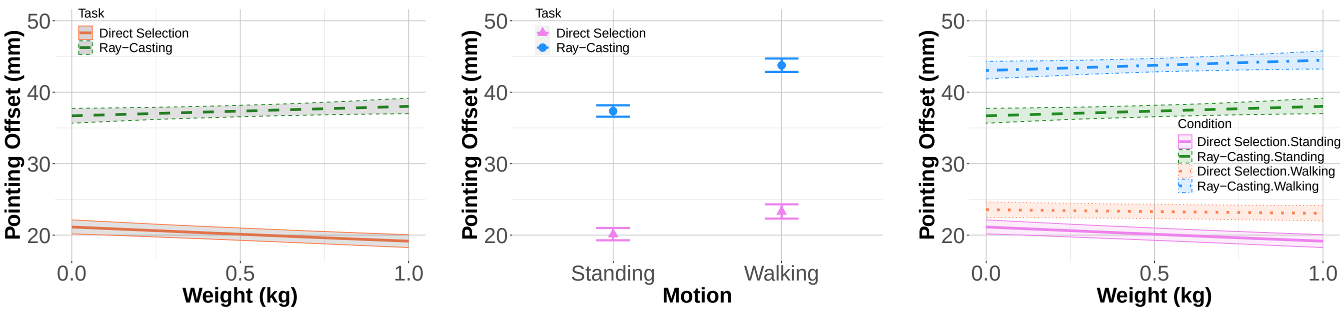
Parameter	Direct Selection		Ray-Casting	
	Estimate (Error)	95% CI	Estimate (Error)	95% CI
Fixed Effects (Independent Variables)				
Intercept	6.06 (0.09)	[5.87, 6.25]	6.74 (0.08)	[6.58, 6.89]
Weight	0.41 (0.05)	[0.31, 0.51]	0.27 (0.07)	[0.14, 0.41]
Walking	0.17 (0.05)	[0.08, 0.27]	0.09 (0.06)	[-0.03, 0.22]
Weight:Walking	-0.07 (0.07)	[-0.21, 0.07]	0.22 (0.10)	[0.02, 0.41]
ID:Weight	-0.11 (0.06)	[-0.24, 0.02]	-0.00 (0.09)	[-0.17, 0.17]
ID:Walking	0.07 (0.06)	[-0.05, 0.19]	0.55 (0.08)	[0.39, 0.71]
ID:Weight:Walking	0.14 (0.09)	[-0.04, 0.31]	-0.24 (0.12)	[-0.48, 0.00]

(a) Effect of Weight on Movement Time **(b) Effect of Motion on Movement Time** **(c) Interaction Effect of Weight and Motion on Movement Time**

Figure 6: Model posterior predictions for movement time across different conditions of weight (0 kg, 0.5 kg, and 1.0 kg) and motion (standing still and walking) under two different tasks (direct selection and ray-casting). Scores correspond to the movement time in milliseconds (higher is worse). The upper bound and lower bound indicate the true value of the estimation lies within the 95% CI. The dot and center line are the predicted median movement time.

Table 4: Summary of the pointing offset model: offset ~ 1 + (1|participant) + order + exercise + width · (weight · movement). We provide the posterior means of parameter estimates (Estimate), posterior error of these estimates (Error), and the upper and lower bound of the 95% CI. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00. The complete set of parameter estimates is in Table 11 in the Appendix B.

Parameter	Direct Selection		Ray-Casting	
	Estimate (Error)	95% CI	Estimate (Error)	95% CI
Fixed Effects (Independent Variables)				
Intercept	2.43 (0.06)	[2.31, 2.54]	2.69 (0.05)	[2.60, 2.79]
Weight	-0.25 (0.07)	[-0.39, -0.11]	-0.00 (0.07)	[-0.14, 0.13]
Walking	0.04 (0.06)	[-0.06, 0.17]	0.12 (0.06)	[-0.01, 0.24]
Weight:Walking	0.11 (0.10)	[-0.09, 0.31]	-0.04 (0.10)	[-0.23, 0.16]
Width:Weight	0.00 (0.00)	[-0.00, 0.00]	0.00 (0.00)	[-0.00, 0.00]
Width:Walking	0.00 (0.00)	[-0.00, 0.00]	0.00 (0.00)	[-0.00, 0.00]
Width:Weight:Walking	-0.00 (0.00)	[-0.00, 0.00]	0.00 (0.00)	[-0.00, 0.00]



(a) Effect of Weight on Pointing Offset **(b) Effect of Motion on Pointing Offset** **(c) Interaction Effect of Weight and Motion on Pointing Offset**

Figure 7: Model posterior predictions for pointing offset across different tasks (direct selection and ray-casting) of weight (0 kg, 0.5 kg, and 1.0 kg) and motion (standing still and walking). Scores correspond to the pointing offset (higher is worse). The upper bound and lower bound indicate the true value of the estimation lies within the 95% CI. The dot and the center line are the predicted median pointing offset.

of 0 indicates no support for weight leading to larger pointing offset in ray-cast. In contrast, the coefficients for motion (0.04, SD = 0.06, CI = [-0.06, 0.17]) under direct selection and the coefficients for motion (0.12, SD = 0.06, CI = [-0.01, 0.24]) under selection via ray-cast are positive, indicating an increase in the pointing offset. The results also show that the effect of walking has a 74% probability of leading to larger pointing offset in direct selection and a 96% probability of leading to larger pointing offset in ray-cast. The Bayes Factor of 1.6×10^{33} in direct selection and 2.6×10^{38} in ray-cast suggest extreme support for walking leading to larger offset size. Additionally, there is a 86% probability of having an interaction effect of weight and motion in direct selection. However, there is only a 36% probability of having an interaction effect of weight and motion in ray-casting.

Figure 7 illustrates the posterior means of parameter estimates and the corresponding bounds of the 95% CI for weight and motion. We can see that the effect of walking substantially increase the pointing offset of the user's end-point location. In summary, our model suggests that, on average, walking increased offset size in both direct selection and ray-cast tasks.

4.1.3 Error Rate. We use Bayesian regression with Poisson distribution to model the effect of encumbrance and walking on the error rate.

Table 5 reports the error rate of the posterior means of parameter estimates, the errors of these estimates, and the upper and lower bounds of the 95% CI. The model suggests that the effect of weight has only a 77% probability of leading to a higher error rate in direct selection and only a 23% probability of leading to a higher error rate in ray-cast, which indicates a moderate effect on the error rate. The model also suggests the effect of walking has a 95% probability of leading to a higher error rate in direct selection and a 100% probability of leading to a higher error rate in ray-casting, which indicates a strong effect of increasing the error rate. The Bayes Factor of 503 in the direct selection and 79 in ray-casting suggest weight has a substantial effect on the error rate. The Bayes Factor of 1302 in direct selection and 9.7×10^{53} in ray-casting suggest walking has an extreme effect on the error rate. It shows that the coefficients for motion (mean = 0.17, error = 0.05, CI = [0.08, 0.27]) under direct selection and motion (0.09, SD = 0.06, CI = [-0.03, 0.22]) under ray-cast are positive, which indicates a slight increase in error rate. In contrast, the coefficients for the weight (0.41, SD = 0.05, CI = [0.31, 0.51]) under direct selection and weight (0.27, SD = 0.07, CI = [0.14, 0.41]) under ray-cast are positive, which indicates a slight increase in error rate. There is no interaction effect of weight and motion under direct selection (-0.07, SD = 0.07, CI = [-0.21, 0.07]), with the probability of 15%. However, there is a positive interaction effect of weight and motion under ray-cast (0.22, SD = 0.10, CI = [0.02, 0.41]), with the probability of 75%.

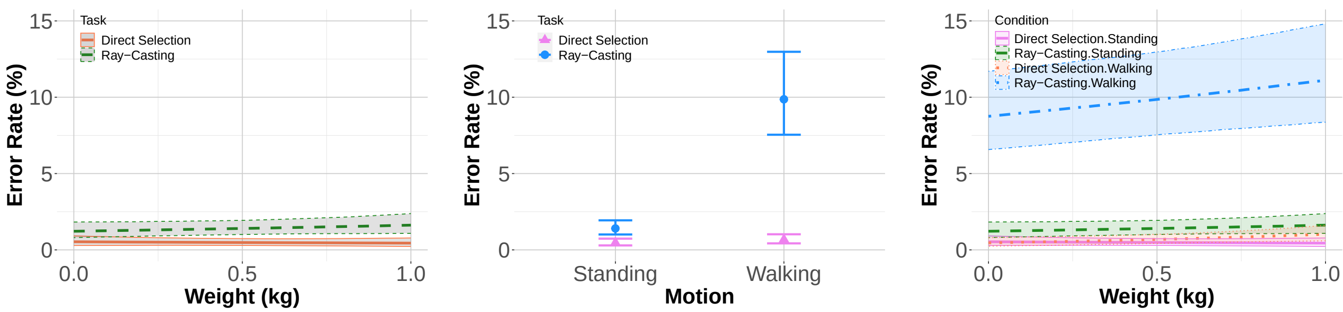
Figure 8 illustrates the posterior means of parameter estimates and the corresponding bounds of the 95% CI for weight and motion. Fitts' law assumes a 4% error rate according to its information theory basis [37, 71]. Thus, the error for direct selection under both standing still and walking conditions remain acceptable, and the error for ray-cast under standing still is acceptable. However, the error rate using ray-cast under walking condition is larger than 4%, which indicates a strong impact on error rate. In summary, our model suggests that, on average, walking and selection via ray-casting result in an increase in error rate, and the error rate under all other conditions remain below acceptable thresholds.

4.1.4 Throughput. To analyse throughput, we kept all data from target acquisition tasks as the metric can evaluate both speed and accuracy. The throughput follows a shifted Log-normal distribution [49, 54]. We used Bayesian regression to model the effect of encumbrance and walking on throughput in a log scale.

Table 6 reports the posterior means of parameter estimates, the errors of these estimates, and the upper and lower bounds of the 95% CI. The model suggests the effect of weight has 0% probability of leading to higher throughput in both direct selection and ray-casting, and the effect of motion has 0% probability of leading to higher throughput in direct selection and 0% probability of leading to higher throughput in ray-casting, which all indicate a strong negative effect on the throughput. The Bayes Factor of 4.6×10^{308} on direct selection and 3.5×10^{38} on ray-casting suggest extreme support for weight leading to lower throughput. Besides, the Bayes

Table 5: Summary of the error rate model: error ~ 1 + (1|participant) + order + exercise + ID · (weight · movement). We provide the posterior means of parameter estimates (Estimate), posterior error of these estimates (Error), and the upper and lower bound of the 95% CI. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00. The complete set of parameter estimates is in Table 12 in the Appendix B.

Parameter	Direct Selection		Ray-Casting	
	Estimate (Error)	95% CI	Estimate (Error)	95% CI
Fixed Effects (Independent Variables)				
Intercept	-5.90 (0.94)	[-7.77, -4.09]	-4.01 (0.59)	[-5.15, -2.87]
Weight	1.02 (1.36)	[-1.64, 3.70]	-0.65 (0.81)	[-2.23, 0.91]
Walking	1.93 (1.22)	[-0.45, 4.39]	2.05 (0.59)	[0.92, 3.20]
Weight:Walking	-1.87 (1.80)	[-5.39, 1.67]	0.58 (0.86)	[-1.09, 2.30]
ID:Weight	-1.57 (1.77)	[-5.12, 1.94]	1.21 (1.02)	[-0.74, 3.20]
ID:Walking	-2.69 (1.62)	[-5.93, 0.44]	-0.11 (0.76)	[-1.55, 1.41]
ID:Weight:Walking	3.73 (2.39)	[-1.06, 8.43]	-0.82 (1.08)	[-2.97, 1.23]



(a) Effect of Weight on Error Rate

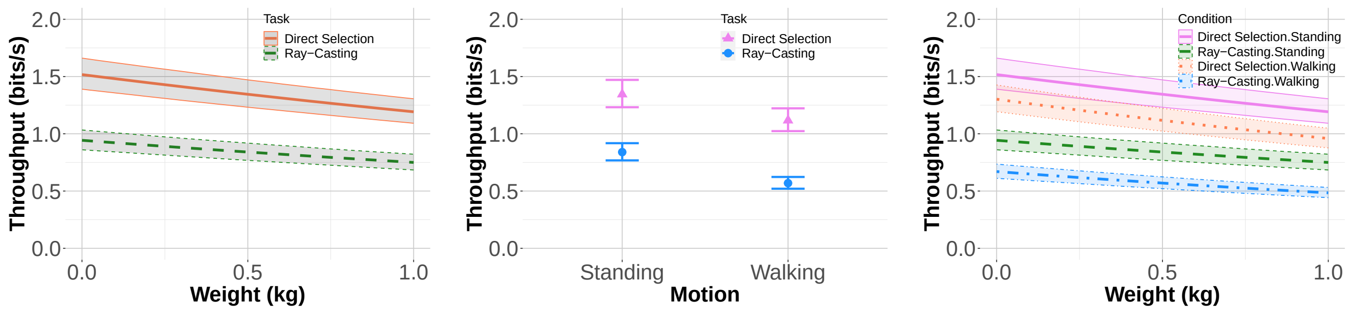
(b) Effect of Motion on Error Rate

(c) Interaction Effect of Weight and Motion on Error Rate

Figure 8: Model posterior predictions for error rate across different conditions of weight (0 kg, 0.5 kg, and 1.0 kg), motion (standing still and walking), and task (selection and ray-cast). Scores correspond to the error rate in percentage (higher is worse). The upper bound and lower bound indicate the true value of the estimation lies within the 95% CI. The dot and the center line are the predicted median error rate.

Table 6: Summary of the target acquisition throughput model: TP ~ 1 + (1|participant) + order + exercise + ID_e · (weight · movement). We provide the posterior means of parameter estimates (Estimate), the posterior error of these estimates (Error), and the upper and lower bound of the 95% CI. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00. The complete set of parameter estimates is in Table 13 in the Appendix B.

Parameter	Direct Selection		Ray-Casting	
	Estimate (Error)	95% CI	Estimate (Error)	95% CI
Fixed Effects (Independent Variables)				
Intercept	-0.71 (0.07)	[-0.85, -0.57]	-1.30 (0.08)	[-1.45, -1.14]
Weight	-0.33 (0.04)	[-0.40, -0.26]	-0.31 (0.06)	[-0.43, -0.19]
Walking	-0.26 (0.03)	[-0.32, -0.19]	-0.21 (0.06)	[-0.32, -0.10]
Weight:Walking	0.08 (0.05)	[-0.03, 0.18]	-0.16 (0.09)	[-0.33, 0.01]
ID_e:Weight	0.10 (0.04)	[0.02, 0.17]	0.09 (0.06)	[-0.03, 0.20]
ID_e:Walking	0.11 (0.03)	[0.05, 0.18]	-0.13 (0.05)	[-0.24, -0.03]
ID_e:Weight:Walking	-0.15 (0.05)	[-0.26, -0.05]	0.06 (0.08)	[-0.10, 0.23]



(a) Effect of Weight on Throughput on Target Acquisition Task

(b) Effect of Motion on Throughput on Target Acquisition Task

(c) Interaction Effect of Weight and Motion on Throughput

Figure 9: Model posterior predictions for throughput across different tasks (direct selection and ray-casting) of weight (0 kg, 0.5 kg, and 1.0 kg) and motion (standing still and walking). Scores correspond to the throughput in bits/s (lower is worse). The upper bound and lower bound indicate the true value of the estimation lies within the 95% CI. The dot and the centre line are the predicted median throughput.

Factor of 1.0×10^{103} on direct selection and 2.1×10^{101} on ray-casting suggest extreme support for walking leading to lower throughput as well. The coefficients for weight (mean = -0.33, error = 0.04, CI = [-0.40, -0.26]) and motion (mean = -0.26, error = 0.03, CI = [-0.32, -0.19]) are negative for direct selection, which indicates a decrease in throughput. In addition, the coefficients for weight (mean = -0.31, error = 0.06, CI = [-0.43, -0.19]) and motion (mean = -0.21, error = 0.06, CI = [-0.32, -0.10]) are negative for ray-casting as well, which indicates a decrease in throughput. The interaction effect of weight and motion on direct selection (mean = 0.08, error = 0.05, CI = [-0.03, 0.18]) and on ray-casting (mean = -0.16, error = 0.09, CI = [-0.33, 0.01]) are not substantial.

Figure 9 illustrates the posterior means of parameter estimates and the corresponding bounds of the 95% CI for weight and motion. We can see that all factors decrease the throughput, with direct selection under weight condition having the highest effect on throughput. In summary, our model suggests that, on average, heavy weight and walking add towards decrease in throughput.

4.1.5 General Findings. 25 out of 30 participants felt that being encumbered affected their performance, mostly linking their performance decline to being fatigued from wearing weights, especially when completing target acquisition via direct selection, e.g., *"I felt tired when wearing weight so I tended to move slower"* (P19). The video recordings also showed that the participants tended to bend their elbows and placed their wrists closer to their upper body and kept their hands lower to reduce the effect of weight. Furthermore, participants also reported that the bigger weight was more tiring than the smaller weight; however, two participants mentioned that the weight did not affect their error rate, e.g., *"I do not feel that the weight will affect the error"* (P27).

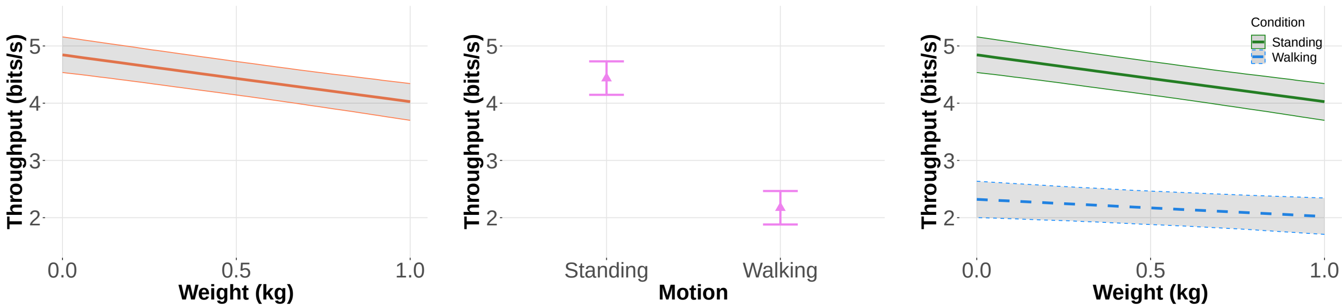
Besides, majority of participants (N = 24) [17, 56] claimed that walking affected their performance in MR tasks, mostly linking the effect to difficulty of aiming at the target, especially when completing target acquisition via ray-casting, e.g., *"It is hard to press the small button [while walking]"* (P8), and *"It is hard to aim the button using ray-casting method [while walking]"* (P32). Some participants (N = 9) [17, 56] noted that it was particularly hard to complete the target acquisition via ray-casting while walking as it requires more effort as compared to while being encumbered, e.g., *"The ray-cast task needed very accurate selection while walking."* (P9). The video recordings also showed that participants tended to briefly stop or moved slowly while walking to aim at the target.

4.2 Text Entry

Each participant completed 30 sentences during the study. We collected a total of 900 sentences during the experiment from all participants. While cleaning the data, we found that 15 sentences were left empty because participants unintentionally pressed the enter button and skipped to the next sentence. We removed these completely empty sentences (1.6% of the dataset) from our data as they do not quantify performance. We evaluate the user performance in the text entry task using throughput.

4.2.1 Throughput. We employ Bayesian regression with a Gaussian distribution to model the effects of encumbrance and walking on the throughput of typing a sentence. This model represents the number of words or characters typed per minute while accounting for the error rate on a normal scale.

Table 7 reports the throughput of the posterior means of parameter estimates, the errors of these estimates, and the upper and lower bounds of the 95% CI in text entry task. The coefficients for motion (mean = -2.52, error = 0.11, CI = [-2.74, -2.31]) and weight (mean = -0.82, error = 0.12, CI = [-1.05, -0.58]) are negative, which indicates a decrease in throughput. The model suggests that both the effect of weight and walking have a 0% probability of leading to higher throughput, which indicates a strong negative effect on the throughput. The Bayes Factor of 4.8695×10^{9} suggests extreme support for weight leading to lower throughput, and the Bayes Factor of 3.2249×10^{103} suggests extreme support for walking leading to lower throughput. There is a 100% probability of having an interaction effect of weight and motion. Figure 8 illustrates the posterior means of parameter estimates and the corresponding bounds of the 95% CI. We can see that both weight and walking have a strong impact on throughput, especially walking has a bigger effect as compared to weight.



(a) Effect of Weight on Throughput (b) Effect of Motion on Throughput (c) Interaction Effect of Weight and Motion on Throughput

Figure 10: Model posterior predictions for throughput in text entry task across different conditions of weight (0 kg, 0.5 kg, and 1.0 kg), and motion (standing still and walking). Scores correspond to the throughput in bits per second (lower is worse). The upper bound and lower bound indicate the true value of the estimation lies within the 95% CI. The dot and the centre line are the predicted median throughput.

Table 7: Summary of the throughput model: TP ~ 1 + (1|participant) + order + exercise + weight · movement. We provide the posterior means of parameter estimates (Estimate), posterior error of these estimates (Error), and the upper and lower bound of their 95% CI. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00. The complete set of parameter estimates is in Table 14 in the Appendix B.

Parameter	Estimate (Error)	95% CI
Fixed Effects (Independent Variables)		
Intercept	4.44 (0.24)	[3.99, 4.92]
Weight	-0.82 (0.12)	[-1.05, -0.58]
Walking	-2.52 (0.11)	[-2.74, -2.31]
Weight:Walking	0.52 (0.17)	[0.19, 0.86]

4.2.2 UER and CER. We further evaluate uncorrected error rate (UER) and corrected error rate (CER) for text entry task. The UER and CER follow a zero-inflated Beta distribution. We used Bayesian regression to model the effect of encumbrance and walking on UER and CER.

Table 8 reports the posterior means of parameter estimates, the errors of these estimates, and the upper and lower bounds of the 95% compatibility/credible interval. The model suggests the effect of weight has a 97% probability of leading to higher UER, and a 24% probability of leading to higher CER, which indicates a decrease on CER. The Bayes Factor of 0.1119 and 0.00476 suggests anecdotal evidence for weight leading to higher UER and CER. The coefficients for weight (mean = 0.55, error = 0.28, CI = [-0.01, 0.11]) indicates slightly increase on UER, and the coefficients for weight (mean = -0.07, error = 0.10, CI = [-0.26, 0.13]) indicates slightly decrease on CER.

When focusing on the motion, the effect of motion has 100% probability of leading to higher UER, which indicates a strong effect on UER, and a 100% probability of leading to higher CER, which indicates a strong effect on CER. The Bayes Factor of 1.32309 suggests anecdotal evidence of leading to higher UER, and the Bayes Factor of 1.0157×10^{22} suggests extreme support for walking leading to higher CER. The coefficients for walking (mean = 0.63, error = 0.22, CI = [0.19, 1.08]) indicates a substantial increase on UER, and the coefficients for walking (mean = 0.47, error = 0.08, CI = [0.32, 0.63]) indicates a strong increase on CER as well.

Figure 11 illustrates the posterior means of parameter estimates and the corresponding bounds of the 95% CI for weight and motion. We can see that weight does not have much effect on CER, but increased UER. Walking increases both UER and CER.

4.2.3 General Findings. During the semi-structured interviews, few participants (N = 3) mentioned that the weight affected their performance when completing the text entry task, e.g., *"The weights are definitely heavy for typing"* (P30). Besides, participants (N = 7) felt that walking affected their typing performance, e.g., *"It is easy to type wrong while walking"* (P12), and *"I would not focus on accuracy while walking and typing. It is better to complete the task first"* (P29).

Additionally, the semi-structured interviews revealed participants' concerns regarding their safety when wearing headsets in their daily lives. Several participants (N=10) claimed that wearing the headset while typing was not safe. For example one participant mentioned: *"It is not safe as when typing, I need to focus on the sentences, which I would forget to check the surroundings"* (P8).

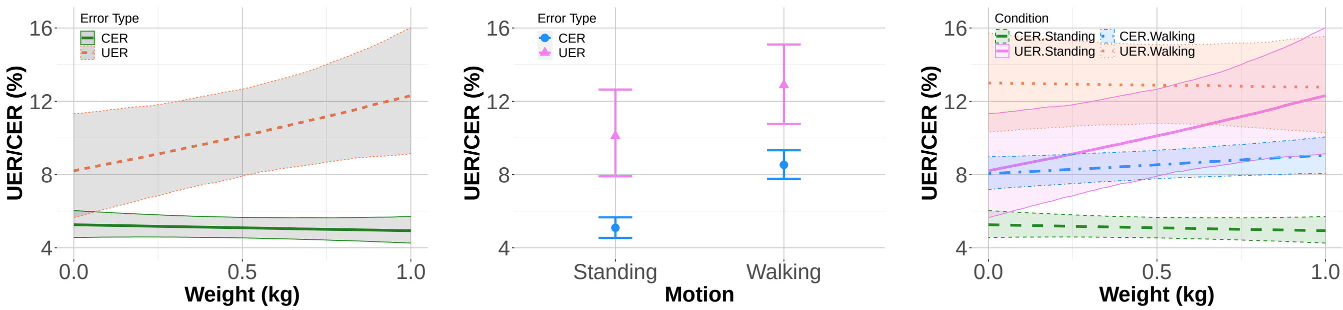
Furthermore, majority of the participants (N = 22) [17, 56] showed concerns about the headset being widely adopted in the society. Specifically, these participants mentioned that they would be hesitant to wear the headset daily unless it became a common practice within their social environment. Some participants claimed that: *"I need to consider how others think about the headset"* (P28), *"If most people are wearing it, I will wear it"* (P20, P25). These findings highlight a broader concern about social acceptance and potential safety issues associated with wearing such technology in everyday settings that need to be taken into account.

5 Discussion

Our findings suggest that both encumbrance and walking substantially affect user performance during target acquisition via direct

Table 8: Summary of the uncorrected error rate and corrected error rate model: UER/CER ~ 1 + (1|participant) + order + exercise + weight · movement. We provide the posterior means of parameter estimates (Estimate), posterior error of these estimates (Error), and the upper and lower bound of their 95% CI. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00. The complete set of parameter estimates is in Table 15 in the Appendix B.

Parameter	Uncorrected Error Rate		Corrected Error Rate	
	Estimate (Error)	95% CI	Estimate (Error)	95% CI
Fixed Effects (Independent Variables)				
Intercept	-1.38 (0.27)	[-1.93, -0.85]	-2.45 (0.10)	[-2.64, -2.27]
Weight	0.55 (0.28)	[-0.01, 1.11]	-0.07 (0.10)	[-0.26, 0.13]
Walking	0.63 (0.22)	[0.19, 1.08]	0.47 (0.08)	[0.32, 0.63]
Weight:Walking	-0.57 (0.34)	[-1.23, 0.09]	0.20 (0.12)	[-0.04, 0.44]



(a) Effect of Weight on UER and CER **(b) Effect of Motion on UER and CER** **(c) Interaction Effect of Weight and Motion on UER and CER**

Figure 11: Model posterior predictions for uncorrected error rate and corrected error rate across different weight (0 kg, 0.5 kg, and 1.0 kg) and motion (standing still and walking). Scores correspond to the uncorrected error rate and corrected error rate in percentage (lower is worse). The upper bound and lower bound indicate the true value of the estimation lies within the 95% CI. The dot and the center line are the predicted median accuracy rate.

selection, target acquisition via ray-casting, and text entry tasks, with walking having a greater impact than encumbrance (1.0 kg weight) at this level. In this section, we discuss the detailed effects of encumbrance and walking, and elaborate on potential reasons behind the observed phenomenon.

5.1 Effects of Encumbrance and Walking on MR Interaction

5.1.1 Encumbrance. Through our study, we identified a strong proportional relationship between weight and movement time. Especially, our results confirm that the negative effect of weight on movement time increases as the weight increases. When using direct selection, 1.0 kg weight (movement time = 832 ms) resulted in a 28% increase on movement time as compared to no-weight condition (movement time = 649 ms). When using ray-cast selection, 1.0 kg weight (movement time = 1861 ms) resulted in a 27% increase on movement time as compared to no weight condition (movement time = 1462 ms). Our findings are consistent with previous research that found that participants took significantly longer time to reach the target while being encumbered during mobile interaction [46, 47]. Ng et al. [47] reported that the overall mean movement times increased by 18% when 1.6 kg weight was held when completing tapping tasks on smartphones. Ng et al. [46] reported that the overall mean movement times increased by 9% when 1.6 kg weight was held during tapping tasks on mobile devices in a one-handed mode.

Furthermore, our results demonstrate that in the target acquisition task, the weight did not show a substantial effect on the pointing offset. The offset has decreased by 9% with 1.0 kg weight (offset = 19.15 mm) under direct selection as compared to the no-weight condition (offset = 21.13 mm), and the offset increased by 4% with 1.0 kg weight (offset = 38.02 mm) as compared to no-weight condition (offset = 36.70 mm). Since the width of the ray-casting task is set to be twice that of the direct selection task, the predicted offset for ray-casting is correspondingly larger than that for direct selection. However, from the Bayes Factor we can see that the changes in offset size for both tasks were not so substantial. This indicates the weight did not affect the offset size. The actual effects are all within 2 mm (roughly between 2-5% of the target size) and are almost negligible for MR interaction in real-world scenarios. Thus, the effect is minimal enough not to impact performance. However, prior research showed that encumbrance with 1.6 kg weight caused a 40% increase in the offset size during smartphone tapping tasks [46]. This misalignment with prior research might be due to the difference in the nature of mobile interaction. Mobile interaction, unlike MR interaction, requires precise fine motor movements as users operate with significantly smaller targets. Encumbrance significantly affects precision in fine motor movements, leading to a more profound effect in touch accuracy [46, 47].

Similarly, our findings did not reveal a strong effect of encumbrance on the error rate. The error rate has decreased 0.1% when encumbered with 1.0 kg weight (error rate = 0.4%) when performing direct selection tasks as compared to no-weight condition (error rate = 0.5%). During the selection via ray-casting, the error rate increased 0.4% when being encumbered with 1.0 kg weight (error rate = 1.6%) as compared to the no-weight condition (error rate = 1.2%). The observed error rates remain acceptable, as Fitts' law assumes a 4% error rate according to its information theory basis [37, 71]. These findings are different from the ones reported in prior research on mobile interaction, which showed the mean error rate increased by 12% when 1.6 kg bag was held during tapping tasks in two-handed interaction mode [47], and the mean error rate increased by 10% when holding 1.6 kg bag and using one-handed interaction mode [46]. This discrepancy with prior research could again be due to the nature of MR interaction as it offers a larger interaction area with greater tolerance for errors, unlike mobile interaction.

Through our study, we identified a strong relationship between weight and target acquisition throughput. When using direct selection, 1.0 kg weight (throughput = 1.19 bits/s) resulted in a 22% decrease as compared to no-weight condition (throughput = 1.52 bits/s). When using ray-cast selection, 1.0 kg weight (throughput = 0.75 bits/s) resulted in a 20% decrease on throughput as compared to the no-weight condition (throughput = 0.94 bits/s). Our findings are inline with previous research on smartphone interaction. Ng et al. [47] reported that the target acquisition throughput increased by 23% when completing tapping tasks on smartphones while 1.6 kg weight was held. This demonstrates that encumbrance significantly impairs throughput in target acquisition tasks both in mobile and MR interaction.

Furthermore, in the text entry task, the model (Figure 10) indicates a negative correlation between weight and throughput. The typing throughput shows a 17% decrease when being encumbered with 1.0 kg weight (throughput = 4.03 bits/s) as compared to the no-weight condition (throughput = 4.84 bits/s). Our findings are similar to previous research on mobile interaction [50], where participants found it is harder to type while being encumbered with different objects.

Finally, we identified that weight does not affect CER but increase UER. Our results indicate that 1.0 kg weight (UER = 12.30%) resulted in a 50% increase in UER as compared to no-weight condition (UER = 8.20%). However, the CER of 1.0 kg weight (CER = 4.90%) is roughly the same as compared to no-weight condition (CER = 5.26%). These findings suggest that, when encumbered, participants chose not to correct errors and left them wrong.

Our quantitative results were reinforced by qualitative insights gathered from semi-structured interviews. Most of the participants (N = 25) [17, 56] reported that being encumbered negatively impacted their performance, often attributing this decline to fatigue caused by wearing weights. Additionally, during the semi-structured interviews, few participants (N = 3) specifically mentioned that the weights hindered their performance during the text entry task.

5.1.2 Walking. Our results indicate a strong proportional relationship between walking and movement time. When using direct selection, walking (movement time = 887 ms) resulted in a 21% increase on movement time as compared to standing still (movement time = 733 ms). When using ray-cast selection, walking (movement time = 2686 ms) resulted in a 63% increase in movement time as compared to standing still (movement time = 1648 ms). This is consistent with prior research, which reported an increase in movement time when completing selection tasks on smartphones (by 31%) while walking [64].

Furthermore, our results demonstrate that walking substantially affected the pointing offset. Walking increased 16% of the offset in direct selection (offset = 23.29 mm) and by 17% in ray-cast selection (offset = 43.76 mm) as compared to standing still (offset = 20.12 mm and offset = 37.35 mm respectively). Ng et al. [46] also reported that using two-handed index finger input posture decreased offset by an average of 16% while carrying bags and walking. Our results support findings by Ng et al. [46] and demonstrate that walking impairs performance in target selection tasks in MR.

Similarly, our results demonstrate that walking using ray-cast also substantially affected the error rate. The error rate increased 0.2% in the direct selection (error rate = 0.7%) and 8.4% in the ray-cast selection (error rate = 9.8%) when walking. Our findings are in line with previous research showing that the error rate has increased by 7% when walking around an obstacle path as compared to being seated [64]. Given that Fitts' law accepts 4% error rate, the value for error rate in direct selection while walking remains acceptable. This observation is similar to the effect of encumbrance on target acquisition. However, the value for error rate in selection via ray-casting was higher than the acceptable threshold, implying that the effect of walking was more severe on selection via ray-casting, than the effect of encumbrance on both target acquisition tasks.

Our results also indicate a strong relationship between walking and target acquisition throughput. When using direct selection, walking (throughput = 1.12 bits/s) resulted in a 16% decrease on throughput as compared to standing still (throughput = 1.34 bits/s). When using ray-cast selection, walking (throughput = 0.57 bits/s) resulted in a 32% increase in throughput as compared to standing still (throughput = 0.84 bits/s).

Furthermore, in the text entry task, the Bayesian model (Figure 10) indicates a strong negative correlation between walking and throughput. The throughput has decreased by 2.26 bits/s (51%) compared to standing still. Our findings are in line with prior research that demonstrated that text entry performance was better while being seated as compared to walking [18]. Our findings are consistent with previous research on mobile interaction [43], where the input speed was significantly lower, and the error rate was significantly higher while walking.

Finally, we identified that walking resulted in increase in both UER and CER. Our results indicate that walking (UER = 12.88%, CER = 8.53%) resulted in a 28% and 68% increase in UER and CER separately as compared to standing still (UER = 10.10%, CER = 5.09%). These findings suggest that the total number of errors during the walking condition was too high, prompting participants to correct some of them to ensure the sentence remained fluent. This is also inline with our qualitative findings from Section 4.2.3, where P29 said *"I would not focus on accuracy while walking and typing. It is better to complete the task first"* (P29), which indicates that the error rate was higher than standing still condition.

Moreover, our quantitative results were supported by qualitative data gathered from semi-structured interviews. The majority of participants (N = 24) reported that walking adversely impacted their performance in MR tasks, primarily attributing this to the difficulty of aiming at the target. Some participants (N = 9) specifically mentioned that completing target acquisition via ray-casting was particularly challenging while walking, as it demands greater precision compared to being encumbered. This finding aligns with prior research, which shows that participants tend to reduce their walking speed to better sample the environment and allocate cognitive resources to other tasks [51]. Furthermore, as P29 mentioned in Section 4.2.3, participants prioritize task completion over accuracy while walking and typing, resulting in a higher error rate, and thus lower text entry throughput.

5.1.3 Encumbrance vs. Walking. Our results demonstrate that in target acquisition and text entry tasks, the effects of walking were more profound as compared to the effects of encumbrance. For example, the movement time in target acquisition via ray-casting while carrying 1.0 kg weight has increased by 27%; while under walking, it has increased by 63%. Majority of our participants (N=22) [17, 56] also stated that the effect of walking was more substantial compared to the effect of encumbrance when completing MR tasks. Several participants (N=7) mentioned that walking affected their performance due to the scene constantly moving, and they were required to focus harder on the task and split their attention to monitor the surrounding environment. Furthermore, our results showed the effect of the weight on participants' performance varied across individuals. This discrepancy can be attributed to the participants' strength levels, which can be influenced by factors like regular physical activity [41, 73]. Therefore, these variations suggest that the effects of encumbrance in MR interaction can be perceived differently depending on one's physical strength and conditioning.

Some participants (N = 9) noted that it was particularly hard to complete the target acquisition via ray-casting while walking as it requires more accuracy as compared to while being encumbered. This observation might have been caused by the nature of the task, as walking not only causes a divide in attention but also induces hand tremors, hence not allowing participants to select targets accurately while walking.

Furthermore, selecting targets using ray-cast requires smaller wrist amplitude for the gesture, thus participants would not feel as tired while encumbered. This is because of the nature of the ray-casting task. It requires movements of an arm within a small area, unlike direct selection, which requires moving the arm at a greater amplitude. Ray-casting allows the elbow to stay closer to the body, thus reducing the weight felt by the shoulder muscles [23, 32].

5.1.4 Interaction Effect of Encumbrance and Walking. Our results demonstrate that in the target acquisition task, the interaction effect of encumbrance and walking does not have a substantial impact on pointing offset, error rate, and throughput. However, there is a substantial interaction effect of weight and motion on movement time. When using ray-casting, encumbered with 1.0 kg weight and walking (movement time = 3097 ms) resulted in a 112% increase in movement time compared to no-weight and standing still condition (movement time = 1462 ms). Furthermore, the interaction effect of weight and motion also affects the throughput of the text entry task. Encumbered with 1.0 kg weight and walking (throughput = 2.02 bits/s) resulted in a 58% decrease in throughput compared to no-weight and standing still conditions (throughput = 4.84 bits/s). These are more profound than the single effect of either being encumbered with a 1.0 kg weight or walking, which suggests that the combined influence of weight and motion amplifies the challenges, making it more difficult for users to perform tasks in MR environments.

5.2 Contrasting the Effects of Encumbrance and Walking

Our findings demonstrate that being encumbered strongly affected movement time during target acquisition tasks. The increase in movement time while being encumbered could be due to the fatigue participants experienced in the mid-air interaction. Hincapié-Ramos et al. [23] and Li et al. [32] demonstrated that the level of fatigue experienced by the shoulder muscles depends on how much weight can be handled while maintaining a required contraction. As a result, participants would make shorter and slower movements after sensing localised fatigue [11]. Hence, encumbrance could potentially have a more pronounced effect in MR interaction, as it mostly requires mid-air gestures. Particularly, given that the weights we used in this research are smaller than the ones used by Ng et al. [46], our findings still demonstrate a substantial deteriorating effect of encumbrance on MR interaction. This further highlights how encumbrance poses even greater challenges on MR interaction, as compared to mobile interaction. Therefore, it is important to study the effects of different situational impairments on MR interaction, especially if people want to use MR as seamlessly as they use mobile devices in their daily lives.

Prior research has shown a negative effect of encumbrance on pointing offset [46, 47] and error rates [64] in mobile interaction. In contrast, our findings did not fully align with these results. Our findings revealed a substantial effect of walking on the pointing offset and error rate on target acquisition via ray-casting. Whereas the values for offset in target acquisition for both via direct selection and ray-casting remain under an acceptable threshold while being encumbered. This discrepancy could be caused by the difference in the nature of mobile and MR interactions. Precisely, in MR devices, we have a larger space for interaction, whereas on mobile devices, the interaction space is limited to the screen size [2]. Additionally, the weights we used were lower than the previous studies (1.6 kg) [46, 47], which could also be a reason that lower weight does not show a substantial effect on offset and error rate. Hence, the effects of encumbrance and walking on the error rates and offset are more profound during mobile interaction than on MR interaction.

The increased fatigue experienced by the participants during direct selection and text entry compared to ray-casting may be attributed to the greater movement amplitude required. Hincapié-Ramos et al. [23] demonstrated that the location of clicks can substantially impact fatigue, i.e., clicking near the centre of the 2D plane while keeping the arm in a bent position causes the least amount of fatigue compared to other positions. Our findings are consistent with the literature; for example, right-handed users need to extend their arms to reach targets in the upper left corner during direct selection. These larger hand movements may lead to increased fatigue, particularly when additional weights are involved. In contrast, when using ray-casting, users can maintain a bent elbow and keep their wrists closer to their upper body following the law of the lever and, hence, reducing fatigue.

Our results show that walking had a greater impact on the pointing offset and error rate in the target acquisition task when compared to encumbrance. This is because walking, as a dynamic activity, requires continuous balance and coordination, which can divide attention between navigating in the environment and completing the tasks [80], whereas encumbrance does not require this divide in attention. Furthermore, the movement of the limbs that occurs due to walking causes the hands and arms to be less steady [81], which results in increased pointing offset and higher error rates. This suggests that a one-size-fits-all strategy is inadequate for addressing the challenges posed by encumbrance and walking. Instead, it's essential to tackle these issues individually, recognising that each presents unique demands. For example, greater focus should be placed on how attention resources are allocated during walking, as this factor plays a critical role in MR interaction and affects interaction performance.

5.3 Addressing the Effects of Encumbrance and Walking on MR Interaction

As our results demonstrate, the effects of encumbrance and walking are different on mobile interaction. Similar to previous studies, we avoid a one-size-fits-all approach, as different technology and design features may be more suitable for different environments and tasks [86].

For instance, eye tracking technology presents a viable alternative to gesture-based interactions. Recognized for its quick and efficient input capabilities, eye tracking has shown potential in enhancing object selection tasks [75, 86]. The primary advantage of eye-tracking is that it allows hands-free interaction and facilitates fast and precise selection of virtual objects [86]. Hence, it can potentially reduce the physical strain on hands and allow users to stay engaged with the MR environment while being encumbered. However, eye-tracking should not be used when the user is walking as it will reduce the user's focus. Therefore, user awareness of their physical surroundings, balance, coordination, and visual attention can be compromised, leading to disorientation, accidents or injuries caused by the obstacles in their real environment.

Furthermore, voice input technology enables users to issue commands verbally, eliminating the need for physical gestures. Its primary strength lies in facilitating complex interactions without requiring physical effort [15]. This method can be applied to a variety of tasks, such as navigating menus, selecting objects, controlling the environment, and triggering specific actions. Voice input is particularly advantageous in scenarios where precision and speed are crucial, as it allows users to perform actions quickly and accurately without moving their hands or arms. Similarly, for the situations when the user is walking, voice input can offer a hands-free and eyes-free alternatives that allow users to issue commands and input text verbally, which takes physical safety into consideration.

Moreover, integrating sensors that monitor the user's physical context in real-time enables MR systems to detect situational impairments (e.g., encumbrance or walking), and dynamically adapt the interaction method accordingly [34]. Instead of explicitly identifying specific activities, the system could also infer situational impairments based on the specific user performance, automatically adjusting the input modality to suit the user's current context. For instance, if the system senses a drop in throughput, it could suggest using hands-free interaction methods such as eye tracking or voice commands to enhance user experience.

In real-world scenarios, it is crucial to design MR systems not only for user comfort but also to ensure that interactions are perceived as natural [27]. Feedback from 6 out of 30 participants indicated feeling awkward when wearing the headset. This reflects a concern for social norms and acceptance when using the headset outdoors. Interaction methods should be developed with an emphasis on aligning with social expectations and natural behaviour [24].

Additionally, the user's physical safety should always be the top priority. In the semi-structured interview, P8 mentioned: *"It is not safe as when typing, I need to focus on the sentences, which I would forget to check the surroundings"* (P8). This is in line with previous research, where individuals using cell phones were significantly less likely to notice different obstacles in their vicinity compared to those not using phones [26]. Furthermore, individuals texting while walking exhibited altered postural control [63] and may overlook obstacles or hazards [26], which increased the risk of accidents. This is called "inattentional blindness"—a psychological phenomenon where individuals fail to notice unexpected events in their environment due to diverted attention [35]. The majority of research on situational impairments is conducted in highly controlled surroundings [59, 82], creating ideal conditions for the participants. However, with the current MR headsets, these ideal conditions might mask their performance flaws. For instance, while Microsoft HoloLens generally performs well under ideal conditions, it may have negative effects in actual military settings, potentially compromising soldiers' safety[5]. This highlights the need for situational impairments research on MR interaction to be taken out of ideal conditions to be able to understand and prevent safety risks that these headsets might impose on users when used outside the lab environment.

5.4 Limitations and Future Work

We acknowledge a number of limitations in our study. First, the study settings were strictly controlled. We simulated encumbrance by attaching a weight on the wrist and removed other factors related to the ergonomics of carrying more common items in our daily lives. Furthermore, we controlled the weight attached to participants'

[5]https://www.techspot.com/news/105242-army-wants-microsoft-substantially-lower-price-80000-hololens.html [Accessed: 2024-11-06]

wrists. However, it is possible that in the real-world context, participants can experience carrying heavier items. We also controlled for a walking path to simulate walking. However, in a real-world scenario, people usually walk on more challenging terrains and experience the presence of other passers-by on their way. Nevertheless, we argue it was necessary to control for these variables in order to eliminate any potential confounding effects and avoid causing harm to our participants [59]. Since we know the prominent effects of encumbrance and walking, we hypothesise that these effects will be more pronounced in real-world scenarios (e.g., carry bags in hand, walking in a busy street). Future work could explore the impact of other types of encumbrance, including wearing day-to-day handbags, a heavy backpack, and different weights attached to different body locations. This could provide further insights into how load distribution affects movement dynamics and overall physical performance. Furthermore, future work could utilise standardised instruments to collect data on participants' physical conditions (e.g., Physical Component Summary) to provide quantitative insights into participants' fitness levels.

Second, we limited the tasks presented in this study to target acquisition and text entry. However, we acknowledge that in real-life scenarios, users may perform more complex tasks in MR. Nevertheless, we controlled for the task to understand the effects of encumbrance and walking on basic MR tasks. Now that we know the substantial effects of encumbrance and walking on these basic tasks, we predict that these effects will be more pronounced when completing more complex tasks in MR. Third, we restricted our participants to using only the index finger of their dominant hand to complete the tasks. We argue that controlling their interaction was necessary to draw a fair comparison among different conditions. Fourth, although we tried to position the targets with different diameters and distances to make sure the task covered wide visual areas, the angles between the targets and users' view were still within the field of view of the headsets. Future work could explore targets located far from the user's centre viewing angle and outside of their view [84, 85].

Finally, the use of the Meta Quest 3, equipped with a video see-through display, may influence user experience compared to optical see-through headsets (e.g., Microsoft HoloLens). Unlike optical see-through systems, the video-based display can introduce visual distortions of real-world objects and may exhibit delays in the world camera feed. However, our choice of Meta Quest 3 was based on the fact that the Meta Quest 3 headset is one of the most popular headsets used among the general population[6]. Furthermore, the high resolution of Meta Quest 3 offers a clearer interface during MR interaction. Finally, compared to other advanced video see-through headsets, the Meta Quest 3 is a cost-effective option.

6 Conclusion

In this study, we investigate the effects of encumbrance and walking on MR interaction performance in target acquisition (direct selection and ray-casting) and text entry tasks. We found that being encumbered and walking slowed participants in completing the tasks. Encumbered with 1.0 kg weight, increased selection movement time by 28%, decreased selection throughput by 22%, decreased text entry throughput by 17%, and increase uncorrected error rate by 50%, but does not affect pointing offset and error rate. Walking led to a 63% increase in ray-casting movement time, a 32% decrease in ray-casting throughput, and a 51% reduction in text entry throughput. We also found that both the pointing offset and error rate were higher in walking conditions as compared to being encumbered. The pointing offset increased by 16% using the direct selection method while walking, the pointing offset increased by 17%, and the error rate increased by 8.4% using the ray-cast interaction method while walking. The corrected error rate increased by 68%. The interaction effect of 1.0 kg weight and walking resulted in a 112% increase in ray-casting movement time and a 58% decrease in text entry throughput. We highlight the importance of considering situational impairments in MR interaction and propose potential strategies and directions to mitigate the impacts of situational impairments. Our findings enhance the understanding of the effects of encumbrance and walking on MR interaction and contribute towards accumulating knowledge in SIIDs research, expanding it to MR.

[6]https://www.sciencefocus.com/reviews/meta-quest-3-review [Accessed: 2024-12-01]

Acknowledgments

Dr. Withana is a recipient of an Australian Research Council Discovery Early Career Award (DECRA) - DE200100479 funded by the Australian Government. We thank our participants for their valuable time. We appreciate the members of the AID-LAB for assisting us in various ways.

References

[1] R Harald Baayen, Douglas J Davidson, and Douglas M Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language* 59, 4 (2008), 390–412.

[2] Evan Barba, Blair MacIntyre, and Elizabeth D Mynatt. 2012. Here we are! Where are we? Locating mixed reality in the age of the smartphone. *Proc. IEEE* 100, 4 (2012), 929–936.

[3] Leon Barnard, Ji Soo Yi, Julie A Jacko, and Andrew Sears. 2005. An empirical comparison of use-in-motion evaluation scenarios for mobile computing devices. *International Journal of Human-Computer Studies* 62, 4 (2005), 487–520.

[4] Joanna Bergström, Tor-Salve Dalsgaard, Jason Alexander, and Kasper Hornbæk. 2021. How to evaluate object selection and manipulation in vr? guidelines from 20 years of studies. In *proceedings of the 2021 CHI conference on human factors in computing systems*. 1–20.

[5] Joe Brailsford, Frank Vetere, and Eduardo Velloso. 2024. Exploring the Association between Moral Foundations and Judgements of AI Behaviour. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–15.

[6] Michelle A Brown, Wolfgang Stuerzlinger, and EJ Mendonça Filho. 2020. The performance of un-instrumented in-air pointing. In *Graphics Interface 2014*. AK Peters/CRC Press, 59–66.

[7] Paul-Christian Bürkner. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software* 80 (2017), 1–28.

[8] Kelly Caine. 2016. Local standards for sample size at CHI. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 981–992.

[9] Bob Carpenter, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus A Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. Stan: A probabilistic programming language. *Journal of statistical software* 76 (2017).

[10] Edwin Chau, Jiakun Yu, Cagatay Goncu, and Anusha Withana. 2021. Composite line designs and accuracy measurements for tactile line tracing on touch surfaces. *Proceedings of the ACM on Human-Computer Interaction* 5, ISS (2021), 1–17.

[11] Jeffrey C Cowley, Jonathan B Dingwell, and Deanna H Gates. 2014. Effects of local and widespread muscle fatigue on movement timing. *Experimental brain research* 232 (2014), 3939–3948.

[12] Edwin John Crossman. 1957. *Factors involved in the predator-prey relationship of rainbow trout (Salmo gairdneri Richardson) and redside shiners (Richardsonius balteatus (Richardson)) in Paul Lake, British Columbia*. Ph. D. Dissertation. University of British Columbia.

[13] Tafadzwa Joseph Dube and Ahmed Sabbir Arif. 2019. Text entry in virtual reality: A comprehensive review of the literature. In *Human-Computer Interaction*.

Recognition and Interaction Technologies: Thematic Area, HCI 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings, Part II 21. Springer, 419–437.
[14] Tafadzwa Joseph Dube and Ahmed Sabbir Arif. 2020. Impact of key shape and dimension on text entry in virtual reality. In *Extended abstracts of the 2020 CHI conference on human factors in computing systems*. 1–10.
[15] Faruk Lawal Ibrahim Dutsinma, Debajyoti Pal, Suree Funilkul, and Jonathan H Chan. 2022. A systematic review of voice assistant usability: An ISO 9241–11 approach. *SN computer science* 3, 4 (2022), 267.
[16] Paul M Fitts. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology* 47, 6 (1954), 381.
[17] Joseph L Fleiss, Bruce Levin, and Myunghee Cho Paik. 2013. *Statistical methods for rates and proportions*. john wiley & sons.
[18] Rachel E Funk, Megan L Taylor, Ceith C Creekmur, Christine M Ohlinger, Ronald H Cox, and William P Berg. 2012. Effect of walking speed on typing performance using an active workstation. *Perceptual and motor skills* 115, 1 (2012), 309–318.
[19] Andrew Gelman. 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.
[20] Mayank Goel, Leah Findlater, and Jacob Wobbrock. 2012. WalkType: using accelerometer data to accomodate situational impairments in mobile touch screen text entry. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2687–2696.
[21] Jorge Goncalves, Zhanna Sarsenbayeva, Niels van Berkel, Chu Luo, Simo Hosio, Sirkka Risanen, Hannu Rintamäki, and Vassilis Kostakos. 2017. Tapping task performance on smartphones in cold temperature. *Interacting with Computers* 29, 3 (2017), 355–367.
[22] Florian Heller, Davy Vanacken, Eva Geurts, and Kris Luyten. 2020. Impact of situational impairment on interaction with wearable displays. In *22nd international conference on human-computer interaction with mobile devices and services*. 1–5.
[23] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1063–1072.
[24] Yi-Ta Hsieh, Antti Jylhä, Valeria Orso, Luciano Gamberini, and Giulio Jacucci. 2016. Designing a willing-to-use-in-public hand gestural interaction technique for smart glasses. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 4203–4215.
[25] Jin Huang and Byungjoo Lee. 2019. Modeling error rates in spatiotemporal moving target selection. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–6.
[26] Ira E Hyman Jr, S Matthew Boss, Breanne M Wise, Kira E McKenzie, and Jenna M Caggiano. 2010. Did you see the unicycling clown? Inattentional blindness while walking and talking on a cell phone. *Applied Cognitive Psychology* 24, 5 (2010), 597–607.
[27] Weiwei Jiang, Maximiliane Windl, Benjamin Tag, Zhanna Sarsenbayeva, and Sven Mayer. 2024. An Immersive and Interactive VR Dataset to Elicit Emotions. *IEEE Transactions on Visualization and Computer Graphics* 30, 11 (2024), 7343–7353. https://doi.org/10.1109/TVCG.2024.3456202
[28] Ricardo Jota, Miguel A Nacenta, Joaquim A Jorge, Sheelagh Carpendale, and Saul Greenberg. 2010. A comparison of ray pointing techniques for very large displays.. In *Graphics Interface*, Vol. 2010. 269–276.
[29] Shaun K Kane, Jacob O Wobbrock, and Ian E Smith. 2008. Getting off the treadmill: evaluating walking user interfaces for mobile devices in public spaces. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*. 109–118.
[30] Jesper Kjeldskov and Jan Stage. 2004. New techniques for usability evaluation of mobile systems. *International journal of human-computer studies* 60, 5-6 (2004), 599–620.
[31] Ben Lambert. 2018. A student's guide to Bayesian statistics. (2018).
[32] Yi Li, Robert Crowther, Jim Smiley, Tim Dwyer, Benjamin Tag, Pourang Irani, and Barrett Ens. 2023. Revisiting Consumed Endurance: A NICE Way to Quantify Shoulder Fatigue in Virtual Reality. In *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*. 1–10.
[33] Min Lin, Rich Goldman, Kathleen J Price, Andrew Sears, and Julie Jacko. 2007. How do people tap when walking? An empirical investigation of nomadic data entry. *International journal of human-computer studies* 65, 9 (2007), 759–769.
[34] Xingyu Bruce Liu, Jiahao Nick Li, David Kim, Xiang'Anthony' Chen, and Ruofei Du. 2024. Human I/O: Towards a Unified Approach to Detecting Situational Impairments. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18.
[35] Arien Mack. 2003. Inattentional blindness: Looking without seeing. *Current directions in psychological science* 12, 5 (2003), 180–184.
[36] Bonnie MacKay, David Dearman, Kori Inkpen, and Carolyn Watters. 2005. Walk'n scroll: a comparison of software-based navigation techniques for different levels of mobility. In *Proceedings of the 7th international conference on Human computer interaction with mobile devices & services*. 183–190.
[37] I Scott MacKenzie. 1992. Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction* 7, 1 (1992), 91–139.
[38] I Scott MacKenzie. 2018. Fitts' law. *The wiley handbook of human computer interaction* 1 (2018), 347–370.
[39] I Scott MacKenzie and Poika Isokoski. 2008. Fitts' throughput and the speed-accuracy tradeoff. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1633–1636.
[40] I Scott MacKenzie and R William Soukoreff. 2003. Phrase sets for evaluating text entry techniques. In *CHI'03 extended abstracts on Human factors in computing systems*. 754–755.
[41] Melissa Mazor, Jeannette Q Lee, Anne Peled, Sarah Zerzan, Chetan Irwin, Margaret A Chesney, Katherine Serrurier, Hani Sbitany, Anand Dhruva, Devorah Sacks, et al. 2018. The effect of yoga on arm volume, strength, and range of motion in women at risk for breast cancer-related lymphedema. *The Journal of Alternative and Complementary Medicine* 24, 2 (2018), 154–160.
[42] Richard McElreath. 2018. *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.
[43] Sachi Mizobuchi, Mark Chignell, and David Newton. 2005. Mobile text entry: relationship between walking speed and text input task difficulty. In *Proceedings of the 7th international conference on Human computer interaction with mobile devices & services*. 122–128.
[44] Terhi Mustonen, Maria Olkkonen, and Jukka Hakkinen. 2004. Examining mobile phone text legibility while walking. In *CHI'04 extended abstracts on Human factors in computing systems*. 1243–1246.
[45] Alexander Ng, Stephen A Brewster, and John Williamson. 2013. The impact of encumbrance on mobile interactions. In *Human-Computer Interaction–INTERACT 2013: 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part III 14*. Springer, 92–109.
[46] Alexander Ng, Stephen A Brewster, and John H Williamson. 2014. Investigating the effects of encumbrance on one-and two-handed interactions with mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1981–1990.
[47] Alexander Ng, John Williamson, and Stephen Brewster. 2015. The effects of encumbrance and mobility on touch-based gesture interactions for mobile phones. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 536–546.
[48] Alexander Ng, John H Williamson, and Stephen A Brewster. 2014. Comparing evaluation methods for encumbrance and walking on interaction with touch-screen mobile devices. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*. 23–32.
[49] Karin Nieuwenhuizen and Jean-Bernard Martens. 2016. Advanced modeling of selection and steering data: beyond Fitts' law. *International Journal of Human-Computer Studies* 94 (2016), 35–52.
[50] Antti Oulasvirta and Joanna Bergstrom-Lehtovirta. 2011. Ease of juggling: studying the effects of manual multitasking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3103–3112.
[51] Antti Oulasvirta, Sakari Tamminen, Virpi Roto, and Jaana Kuorelahti. 2005. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 919–928.
[52] Julian Petford, Miguel A Nacenta, and Carl Gutwin. 2018. Pointing all around you: selection performance of mouse and ray-cast pointing in full-coverage displays. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–14.
[53] José Pinheiro and Douglas Bates. 2006. *Mixed-effects models in S and S-PLUS*. Springer science & business media.
[54] Réjean Plamondon and Adel M Alimi. 1997. Speed/accuracy trade-offs in target-directed movements. *Behavioral and brain sciences* 20, 2 (1997), 279–303.
[55] Pamela Ravasio, Sissel Guttormsen Schär, and Helmut Krueger. 2004. In pursuit of desktop evolution: User problems and practices with modern desktop systems. *ACM Transactions on Computer-Human Interaction (TOCHI)* 11, 2 (2004), 156–180.
[56] Robert Rosenthal and Ralph L Rosnow. 2008. *Essentials of behavioral research: Methods and data analysis*.
[57] Riccardo Russo. 2021. *Statistics for the Behavioural Sciences: An Introduction to Frequentist and Bayesian Approaches* (2nd ed.). Routledge.
[58] Zhanna Sarsenbayeva, Jorge Goncalves, Juan García, Simon Klakegg, Sirkka Rissanen, Hannu Rintamäki, Jari Hannu, and Vassilis Kostakos. 2016. Situational impairments to mobile interaction in cold environments. In *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing*. 85–96.
[59] Zhanna Sarsenbayeva, Niels van Berkel, Danula Hettiachchi, Weiwei Jiang, Tilman Dingler, Eduardo Velloso, Vassilis Kostakos, and Jorge Goncalves. 2019. Measuring the effects of stress on mobile interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 1 (2019), 1–18.
[60] Zhanna Sarsenbayeva, Niels van Berkel, Weiwei Jiang, Danula Hettiachchi, Vassilis Kostakos, and Jorge Goncalves. 2019. Effect of Ambient Light on Mobile Interaction. In *Human-Computer Interaction – INTERACT 2019*, David Lamas, Fernando Loizides, Lennart Nacke, Helen Petrie, Marco Winckler, and Panayiotis

Zaphiris (Eds.). Springer International Publishing, Cham, 465–475.
[61] Zhanna Sarsenbayeva, Niels van Berkel, Chu Luo, Vassilis Kostakos, and Jorge Goncalves. 2017. Challenges of situational impairments during interaction with mobile devices. In *Proceedings of the 29th australian conference on computer-human interaction*. 477–481.
[62] Zhanna Sarsenbayeva, Niels van Berkel, Eduardo Velloso, Vassilis Kostakos, and Jorge Goncalves. 2018. Effect of distinct ambient noise types on mobile interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 1–23.
[63] Siobhan M Schabrun, Wolbert van den Hoorn, Alison Moorcroft, Cameron Greenland, and Paul W Hodges. 2014. Texting and walking: strategies for postural control and implications for safety. *PloS one* 9, 1 (2014), e84312.
[64] Bastian Schildbach and Enrico Rukzio. 2010. Investigating selection and reading performance on a mobile phone while walking. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*. 93–102.
[65] Jonas Schjerlund, Kasper Hornbæk, and Joanna Bergström. 2021. Ninja hands: Using many hands to improve target selection in vr. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–14.
[66] Martin Schmettow. 2021. *New statistics for design researchers*. Springer.
[67] Andrew Sears, Min Lin, Julie Jacko, and Yan Xiao. 2003. When computers fade: Pervasive computing and situationally-induced impairments and disabilities. In *HCI international*, Vol. 2. 1298–1302.
[68] Claude Elwood Shannon. 1948. A mathematical theory of communication. *The Bell system technical journal* 27, 3 (1948), 379–423.
[69] Tom AB Snijders and Roel Bosker. 2011. Multilevel analysis: An introduction to basic and advanced multilevel modeling. (2011).
[70] R William Soukoreff and I Scott MacKenzie. 2003. Metrics for text entry research: An evaluation of MSD and KSPC, and a new unified error metric. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 113–120.
[71] R William Soukoreff and I Scott MacKenzie. 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *International journal of human-computer studies* 61, 6 (2004), 751–789.
[72] Marco Speicher, Anna Maria Feit, Pascal Ziegler, and Antonio Krüger. 2018. Selection-based text entry in virtual reality. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–13.
[73] Shirley Telles, Sachin Kr Sharma, Arti Yadav, Nilkamal Singh, and Acharya Balkrishna. 2014. Immediate changes in muscle strength and motor speed following yoga breathing. *Indian J Physiol Pharmacol* 58, 1 (2014), 22–9.
[74] Garreth W. Tigwell, Zhanna Sarsenbayeva, Benjamin M. Gorman, David R. Flatla, Jorge Goncalves, Yeliz Yesilada, and Jacob O. Wobbrock. 2019. Addressing the Challenges of Situationally-Induced Impairments and Disabilities in Mobile Interaction. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3290607.3299029
[75] Jayson Turner, Jason Alexander, Andreas Bulling, and Hans Gellersen. 2015. Gaze+ RST: integrating gaze and multitouch for remote rotate-scale-translate tasks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 4179–4188.
[76] Aki Vehtari, Andrew Gelman, Daniel Simpson, Bob Carpenter, and Paul-Christian Bürkner. 2021. Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with discussion). *Bayesian analysis* 16, 2 (2021), 667–718.
[77] Samangi Wadinambiarachchi, Ryan M Kelly, Saumya Pareek, Qiushi Zhou, and Eduardo Velloso. 2024. The Effects of Generative AI on Design Fixation and Divergent Thinking. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18.
[78] Eric-Jan Wagenmakers, Ruud Wetzels, Denny Borsboom, and Han LJ Van Der Maas. 2011. Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). (2011).
[79] Alan Traviss Welford. 1968. Fundamentals of skill. (1968).
[80] Christopher D Wickens. 2002. Multiple resources and performance prediction. *Theoretical issues in ergonomics science* 3, 2 (2002), 159–177.
[81] Christopher D Wickens, William S Helton, Justin G Hollands, and Simon Banbury. 2021. *Engineering psychology and human performance*. Routledge.
[82] Jacob O Wobbrock. 2019. Situationally-induced impairments and disabilities. *Web Accessibility: A Foundation for Research* (2019), 59–92.
[83] Jacob O Wobbrock, Kristen Shinohara, and Alex Jansen. 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*. 1639–1648.
[84] Difeng Yu, Hai-Ning Liang, Kaixuan Fan, Heng Zhang, Charles Fleming, and Konstantinos Papangelis. 2019. Design and evaluation of visualization techniques of off-screen and occluded targets in virtual reality environments. *IEEE transactions on visualization and computer graphics* 26, 9 (2019), 2762–2774.
[85] 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 Transactions on Graphics (TOG)* 38, 6 (2019), 1–13.
[86] Difeng Yu, Xueshi Lu, Rongkai Shi, Hai-Ning Liang, Tilman Dingler, Eduardo Velloso, and Jorge Goncalves. 2021. Gaze-supported 3d object manipulation in virtual reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.
[87] Difeng Yu, Qiushi Zhou, Benjamin Tag, Tilman Dingler, Eduardo Velloso, and Jorge Goncalves. 2020. Engaging participants during selection studies in virtual reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 500–509.
[88] Jiakun Yu, Supun Kuruppu, Biyon Fernando, Praneeth Bimsara Perera, Yuta Sugiura, Sriram Subramanian, and Anusha Withana. 2024. IrOnTex: Using Ironable 3D Printed Objects to Fabricate and Prototype Customizable Interactive Textiles. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 3 (2024), 1–26.
[89] Jiakun Yu, Praneeth Bimsara Perera, Rahal Viddusha Perera, Mohammad Mirkhalaf, and Anusha Withana. 2024. Fabricating Customizable 3D Printed Pressure Sensors by Tuning Infill Characteristics. *IEEE Sensors Journal* (2024).
[90] Shumin Zhai, Jing Kong, and Xiangshi Ren. 2004. Speed–accuracy tradeoff in Fitts' law tasks—on the equivalency of actual and nominal pointing precision. *International journal of human-computer studies* 61, 6 (2004), 823–856.
[91] Mingrui Ray Zhang, Shumin Zhai, and Jacob O Wobbrock. 2019. Text entry throughput: Towards unifying speed and accuracy in a single performance metric. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.
[92] Qiushi Zhou, Brandon Victor Syiem, Beier Li, Jorge Goncalves, and Eduardo Velloso. 2024. Reflected Reality: Augmented Reality through the Mirror. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 4 (2024), 1–28.
[93] Joanne E Zucco, Bruce H Thomas, and Karen Grimmer. 2006. Evaluation of four wearable computer pointing devices for drag and drop tasks when stationary and walking. In *2006 10th IEEE International Symposium on Wearable Computers*. IEEE, 29–36.

A Text Entry - WPM Results

We further analyses the text entry task using words per minute (WPM). The WPM follows a shifted Log-normal distribution. We used Bayesian regression to model the effect of encumbrance and walking on WPM. The results are inline with our main findings, where 1.0 kg weight resulted in a 16% decrease in WPM, and walking resulted in a 45% decrease in WPM. The interaction effect of 1.0 kg weight and walking resulted in a 54% decrease in WPM.

Table 9 reports the posterior means of parameter estimates, the errors of these estimates, and the upper and lower bounds of the 95% compatibility/credible interval. The model suggests the effect of weight has 0% probability of leading to higher WPM, and the effect of motion has 0% probability of leading to higher WPM, which all indicate a strong negative effect on WPM. The Bayes Factor of 0 suggests anecdotal evidence for weight leading to lower WPM, and the Bayes Factor of 3.2249×10^{103} suggests extreme support for walking leading to lower WPM. The coefficients for weight (mean = -0.18, error = 0.04, CI = [-0.26, -0.09]) and motion (mean = -0.61, error = 0.04, CI = [-0.68, -0.53]) are both negative, which indicates weight and motion decreases the WPM. There is no interaction effect of weight and motion (mean = 0.01, error = 0.06, CI = [-0.11, 0.13]).

Figure 12 illustrates the posterior means of parameter estimates and the corresponding bounds of the 95% CI for weight and motion. We can see that all factors decrease the WPM, with walking having substantial effect on WPM than weight. In summary, our model suggests that, on average, weight and walking add towards decrease in WPM, and walking affects more than weight.

Through our study, we identified a strong relationship between walking and WPM. Our results indicate that 1.0 kg weight (WPM = 12.81) resulted in a 16% decrease in WPM as compared to no-weight condition (WPM = 15.29). Walking (WPM = 7.68) resulted

Table 9: Summary of the WPM model: WPM ~ 1 + (1|participant) + order + exercise + weight · movement. We provide the posterior means of parameter estimates (Estimate), posterior error of these estimates (Error), and the upper and lower bound of their 95% CI. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Estimate (Error)	95% CI
Regression Coefficients		
Intercept	2.48 (0.07)	[2.35, 2.61]
Weight	-0.18 (0.04)	[-0.26, -0.09]
Walking	-0.61 (0.04)	[-0.68, -0.53]
Weight:Walking	0.01 (0.06)	[-0.11, 0.13]
Fixed Effects		
Order	0.07 (0.01)	[0.06, 0.08]
Exercise	0.00 (0.01)	[-0.01, 0.01]
Random Effects		
Participant: sd(Intercept)	0.20 (0.03)	[0.15, 0.27]
Further Distributional Parameters		
sigma	0.36 (0.01)	[0.34, 0.38]
ndt	0.01 (0.01)	[0.00, 0.03]

(a) Effect of Weight on WPM

(b) Effect of Motion on WPM

(c) Interaction Effect of Weight and Motion on WPM

Figure 12: Model posterior predictions for words per minute across different weight (0 kg, 0.5 kg, and 1.0 kg) and motion (standing still and walking). Scores correspond to the words per minute (lower is worse). The upper bound and lower bound indicate the true value of the estimation lies within the 95% CI. The dot and the center line are the predicted median accuracy rate.

in a 45% decrease as compared to standing still (WPM = 14.00). Additionally, walking while encumbered with 1.0 kg weight (WPM = 7.06) resulted in a 54% decrease as compared to standing still with no-weight (WPM = 15.29). This is inline with our throughput results, where the interaction effect of weight and motion substantially affects the performance.

B Tables of the Complete Set of Parameter Estimates

Table 10: Summary of the movement time model: MT ~ 1 + (1|participant) + order + exercise + ID · (weight · movement). We provide the posterior means of parameter estimates (Estimate), posterior error of these estimates (Error), and the upper and lower bound of the 95% CI. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Direct Selection		Ray-Casting	
	Estimate (Error)	95% CI	Estimate (Error)	95% CI
Fixed Effects (Independent Variables)				
Intercept	6.06 (0.09)	[5.87, 6.25]	6.74 (0.08)	[6.58, 6.89]
Weight	0.41 (0.05)	[0.31, 0.51]	0.27 (0.07)	[0.14, 0.41]
Walking	0.17 (0.05)	[0.08, 0.27]	0.09 (0.06)	[-0.03, 0.22]
Weight:Walking	-0.07 (0.07)	[-0.21, 0.07]	0.22 (0.10)	[0.02, 0.41]
ID:Weight	-0.11 (0.06)	[-0.24, 0.02]	-0.00 (0.09)	[-0.17, 0.17]
ID:Walking	0.07 (0.06)	[-0.05, 0.19]	0.55 (0.08)	[0.39, 0.71]
ID:Weight:Walking	0.14 (0.09)	[-0.04, 0.31]	-0.24 (0.12)	[-0.48, 0.00]
Fixed Effects (Covariates)				
Order	-0.05 (0.00)	[-0.06, -0.05]	-0.04 (0.00)	[-0.04, -0.03]
Exercise	-0.01 (0.01)	[-0.03, 0.00]	-0.01 (0.01)	[-0.02, 0.00]
ID	0.28 (0.04)	[0.20, 0.36]	0.48 (0.06)	[0.37, 0.59]
Random Effects				
Participant (SD)	0.31 (0.04)	[0.24, 0.40]	0.24 (0.03)	[0.18, 0.32]
Further Distributional Parameters				
sigma	0.46 (0.00)	[0.45, 0.46]	0.63 (0.00)	[0.62, 0.64]
ndt	177.66 (0.77)	[176.00, 179.02]	175.36 (3.15)	[168.76, 181.18]

Table 11: Summary of the pointing offset model: offset ~ 1 + (1|participant) + order + exercise + width · (weight · movement). We provide the posterior means of parameter estimates (Estimate), posterior error of these estimates (Error), and the upper and lower bound of the 95% CI. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Direct Selection		Ray-Casting	
	Estimate (Error)	95% CI	Estimate (Error)	95% CI
Fixed Effects (Independent Variables)				
Intercept	2.43 (0.06)	[2.31, 2.54]	2.69 (0.05)	[2.60, 2.79]
Weight	-0.25 (0.07)	[-0.39, -0.11]	-0.00 (0.07)	[-0.14, 0.13]
Walking	0.04 (0.06)	[-0.06, 0.17]	0.12 (0.06)	[-0.01, 0.24]
Weight:Walking	0.11 (0.10)	[-0.09, 0.31]	-0.04 (0.10)	[-0.23, 0.16]
Width:Weight	0.00 (0.00)	[-0.00, 0.00]	0.00 (0.00)	[-0.00, 0.00]
Width:Walking	0.00 (0.00)	[-0.00, 0.00]	0.00 (0.00)	[-0.00, 0.00]
Width:Weight:Walking	-0.00 (0.00)	[-0.00, 0.00]	0.00 (0.00)	[-0.00, 0.00]
Fixed Effects (Covariates)				
Order	0.01 (0.00)	[0.00, 0.02]	0.00 (0.00)	[-0.00, 0.00]
Exercise	0.01 (0.00)	[-0.00, 0.01]	0.00 (0.00)	[-0.00, 0.01]
Width	0.01 (0.00)	[0.00, 0.01]	0.01 (0.00)	[0.01, 0.01]
Random Effects				
Participant (SD)	0.11 (0.02)	[0.08, 0.14]	0.04 (0.01)	[0.02, 0.06]
Further Distributional Parameters				
sigma	0.59 (0.00)	[0.58, 0.60]	0.59 (0.00)	[0.58, 0.60]
ndt	0.00 (0.00)	[0.00, 0.02]	0.02 (0.02)	[0.00, 0.07]

Table 12: Summary of the error rate model: error ~ 1 + (1|participant) + order + exercise + ID · (weight · movement). We provide the posterior means of parameter estimates (Estimate), posterior error of these estimates (Error), and the upper and lower bound of the 95% CI. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Direct Selection		Ray-Casting	
	Estimate (Error)	**95% CI**	**Estimate (Error)**	**95% CI**
Fixed Effects (Independent Variables)				
Intercept	-5.90 (0.94)	[-7.77, -4.09]	-4.01 (0.59)	[-5.15, -2.87]
Weight	1.02 (1.36)	[-1.64, 3.70]	-0.65 (0.81)	[-2.23, 0.91]
Walking	1.93 (1.22)	[-0.45, 4.39]	2.05 (0.59)	[0.92, 3.20]
Weight:Walking	-1.87 (1.80)	[-5.39, 1.67]	0.58 (0.86)	[-1.09, 2.30]
ID:Weight	-1.57 (1.77)	[-5.12, 1.94]	1.21 (1.02)	[-0.74, 3.20]
ID:Walking	-2.69 (1.62)	[-5.93, 0.44]	-0.11 (0.76)	[-1.55, 1.41]
ID:Weight:Walking	3.73 (2.39)	[-1.06, 8.43]	-0.82 (1.08)	[-2.97, 1.23]
Fixed Effects (Covariates)				
Order	0.05 (0.06)	[-0.07, 0.17]	0.04 (0.02)	[-0.00, 0.08]
Exercise	0.03 (0.03)	[-0.02, 0.08]	-0.01 (0.02)	[-0.04, 0.03]
ID	0.36 (1.08)	[-1.82, 2.41]	-0.59 (0.71)	[-2.02, 0.75]
Random Effects				
Participant (SD)	0.78 (0.21)	[0.44, 1.25]	0.71 (0.11)	[0.53, 0.96]

Table 13: Summary of the target acquisition throughput model: TP ~ 1 + (1|participant) + order + exercise + ID_e · (weight · movement). We provide the posterior means of parameter estimates (Estimate), the posterior error of these estimates (Error), and the upper and lower bound of the 95% CI. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Direct Selection		Ray-Casting	
	Estimate (Error)	**95% CI**	**Estimate (Error)**	**95% CI**
Fixed Effects (Independent Variables)				
Intercept	-0.71 (0.07)	[-0.85, -0.57]	-1.30 (0.08)	[-1.45, -1.14]
Weight	-0.33 (0.04)	[-0.40, -0.26]	-0.31 (0.06)	[-0.43, -0.19]
Walking	-0.26 (0.03)	[-0.32, -0.19]	-0.21 (0.06)	[-0.32, -0.10]
Weight:Walking	0.08 (0.05)	[-0.03, 0.18]	-0.16 (0.09)	[-0.33, 0.01]
ID_e:Weight	0.10 (0.04)	[0.02, 0.17]	0.09 (0.06)	[-0.03, 0.20]
ID_e:Walking	0.11 (0.03)	[0.05, 0.18]	-0.13 (0.05)	[-0.24, -0.03]
ID_e:Weight:Walking	-0.15 (0.05)	[-0.26, -0.05]	0.06 (0.08)	[-0.10, 0.23]
Fixed Effects (Covariates)				
Order	0.04 (0.00)	[0.03, 0.04]	0.03 (0.00)	[0.02, 0.04]
Exercise	0.01 (0.01)	[-0.00, 0.02]	0.01 (0.01)	[-0.00, 0.02]
ID_e	0.95 (0.02)	[0.90, 0.99]	0.86 (0.04)	[0.79, 0.94]
Random Effects				
Participant (SD)	0.24 (0.03)	[0.19, 0.32]	0.24 (0.04)	[0.18, 0.32]
Further Distributional Parameters				
sigma	0.39 (0.00)	[0.39, 0.40]	0.71 (0.00)	[0.70, 0.72]
ndt	0.00 (0.00)	[0.00, 0.00]	0.00 (0.00)	[0.00, 0.00]

Table 14: Summary of the throughput model: TP ~ 1 + (1|participant) + order + exercise + weight · movement. We provide the posterior means of parameter estimates (Estimate), posterior error of these estimates (Error), and the upper and lower bound of their 95% CI. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Estimate (Error)	95% CI
Fixed Effects (Independent Variables)		
Intercept	4.44 (0.24)	[3.99, 4.92]
Weight	-0.82 (0.12)	[-1.05, -0.58]
Walking	-2.52 (0.11)	[-2.74, -2.31]
Weight:Walking	0.52 (0.17)	[0.19, 0.86]
Fixed Effects (Covariates)		
Order	0.17 (0.02)	[0.13, 0.21]
Exercise	-0.00 (0.02)	[-0.04, 0.04]
Random Effects		
Participant (SD)	0.77 (0.11)	[0.58, 1.03]
Further Distributional Parameters		
sigma	1.05 (0.03)	[1.00, 1.10]

Table 15: Summary of the uncorrected error rate and corrected error rate model: UER/CER ~ 1 + (1|participant) + order + exercise + weight · movement. We provide the posterior means of parameter estimates (Estimate), posterior error of these estimates (Error), and the upper and lower bound of their 95% CI. All parameter estimates converged with an ESS well above 1000 and an R-hat of 1.00.

Parameter	Uncorrected Error Rate		Corrected Error Rate	
	Estimate (Error)	**95% CI**	**Estimate (Error)**	**95% CI**
Regression Coefficients				
Intercept	-1.38 (0.27)	[-1.93, -0.85]	-2.45 (0.10)	[-2.64, -2.27]
Weight	0.55 (0.28)	[-0.01, 1.11]	-0.07 (0.10)	[-0.26, 0.13]
Walking	0.63 (0.22)	[0.19, 1.08]	0.47 (0.08)	[0.32, 0.63]
Weight:Walking	-0.57 (0.34)	[-1.23, 0.09]	0.20 (0.12)	[-0.04, 0.44]
Fixed Effects				
Order	-0.02 (0.04)	[-0.09, 0.05]	-0.03 (0.02)	[-0.06, -0.00]
Exercise	0.01 (0.01)	[-0.01, 0.04]	0.00 (0.01)	[-0.01, 0.01]
Random Effects				
Participant: sd(Intercept)	0.40 (0.10)	[0.23, 0.61]	0.16 (0.04)	[0.09, 0.25]
Further Distributional Parameters				
phi	1.94 (0.14)	[1.67, 2.22]	25.08 (1.53)	[22.20, 28.23]
zi	0.61 (0.02)	[0.57, 0.64]	0.30 (0.02)	[0.27, 0.34]