

Investigating 3D Selection Performance While Running in Extended Reality: an In-depth Study of the Effect of Speed

Ang Li, Jarrod Knibbe, Kiichiro Tatsuzawa, Gianluca Demartini, Stephen Viller, Guy Wallis, Wolfgang Stuerzlinger, Maxime Cordeil

Abstract—We explore how running speed affects target selection in immersive environments. Through a Fitts’ tapping task, we investigate direct touch selections while participants are standing or running (at speeds up to 9 km/h). We conduct a multifaceted controlled study (n=22) to understand the effect of speed, pointing phase, and foot-strike on selection performance. We find that speed dramatically increases movement time, error rate, perceived exertion, and task load. The increased movement time is driven by the correction phase, which doubles in duration at 9 km/h, while the ballistic phase remains stable. Our exploratory analysis further reveals that a higher cadence reduces the error rate at lower speeds (4.5 km/h), while this effect disappears at higher speeds (9 km/h), as participants’ motor control in target selection is significantly disrupted by the movement oscillations. As the selection error rate quickly exceeds 30% when speed increases, even with accurate, fast external tracking, we argue that direct touch interaction is likely inappropriate for running. Further efforts are needed to facilitate selection interaction at higher running speeds.

Index Terms—Running, Extended Reality, Fitts’ Law, Interaction in Motion, Target Selection

I. INTRODUCTION

THERE is a vision of future data-driven sports being supported by immersive, lightweight, heads-up displays [1]. These devices^{1,2} might enable athletes to examine their performance data in real-time, respond to messages, navigate routes, and more. But how should users interact with these devices? How should they navigate between data fields and select key points of interest on maps?

Today’s interactions in data-driven sports take place primarily on sports watches and smartphones [1]–[3]. Watches, traditionally, have offered limited interactions, allowing inputs to be reduced to a small number of button presses or swipes. These interactions, however, can easily alter the arm motions of a given sport and, thus, interrupt flow. Smartphones, conversely, offer much richer interactions, but also require finer inputs. The visual attention and relative precision required cause athletes to slow down or stop to interact [4]. The move to Extended Reality (XR) displays can offer larger screen space

and richer details, but does not easily support direct, on-device interaction. So, how should we interact?

In this context, the options include pointing, gestures, gaze, voice, controllers, or any combination. Conceptually, these approaches all have strengths and weaknesses. Voice control is quickly becoming more powerful through AI, but remains socially awkward [5], and could be interrupted under (breathless) exertion [6]. Controllers might support specific input, but require additional devices, which reduces the inherent benefits of the lightweight glasses [7]. Gestures do not require accuracy or, necessarily, spatial relevance, but they are abstract and do not easily support specific, detailed input [8]. Gaze interaction can be simple and direct, but is inherently a dual-task [9], necessary for both maintaining spatial awareness and any input. These trade-offs become particularly salient during running: runners may need to dismiss an alert, change a data field, select a location on a route, or mark an event without wanting to stop or slow down [4]. For this reason, pointing has emerged as an early contender (e.g., [10]–[12]) – it is simple, direct, specific, and well-understood in interaction design.

Fitts’ law is a predictive model used to estimate how quickly and accurately a person can move to and select a target on a graphical user interface, depending mainly on factors like the size of and the distance between targets [12], [13]. While Fitts’ law has been extensively studied in XR, most of the existing studies have focused on stationary or seated contexts [12]. A few studies have investigated Fitts’ law in mobile scenarios [14]–[18]. Collectively, these works demonstrate the feasibility of pointing under locomotion, but report that movement degrades selection performance. However, all of these works explored only low-intensity movement scenarios up to a walking speed. To our knowledge, there is no research investigating Fitts’ law in more intense movement scenarios, such as running, where body oscillations and gait instability are stronger and may fundamentally change selection dynamics [16]. Moreover, prior research has reported outcomes related to Fitts’ law, such as movement (*MT*) and error rate (*ER*), but lacked contributions towards the understanding of the mechanism that accounts for the reported performance degradation [12]. Besides, not all of these previous studies followed the ISO-9421:411 standard [19] for evaluating the tapping task, making cross-comparison with prior results challenging [14], [15]. Prior studies were also limited in terms of tracking methods, as movement was tracked only through the headset [15]–[18], whose tracking capabilities and accuracy

Ang Li, Jarrod Knibbe, Kiichiro Tatsuzawa, Gianluca Demartini, Stephen Viller, Guy Wallis, and Maxime Cordeil are with the University of Queensland. E-mail: angli21@sigchi.org, viller@acm.org, {j.knibbe | k.tatsuzawa | g.demartini | g.wallis | m.cordeil}@uq.edu.au

Wolfgang Stuerzlinger is with Simon Fraser University. E-mail: w.s@sfu.ca
Manuscript received April 19, 2021; revised August 16, 2021.

¹Engo Eyewear

²ActiveLook

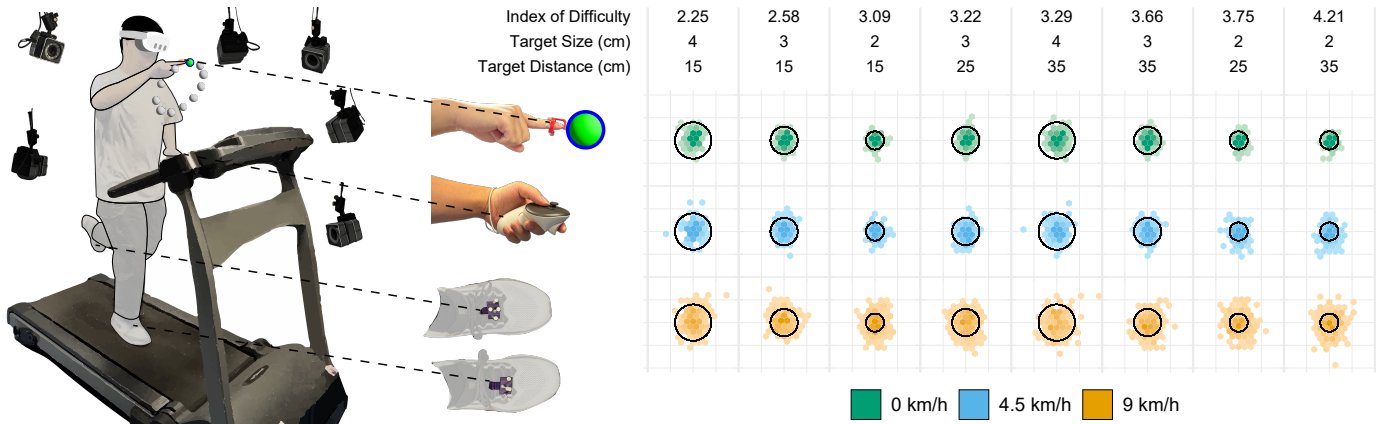


Fig. 1. We conduct a Fitts' law study with direct virtual touch tasks while participants are running or standing on a treadmill at different speeds. We capture participants' fingertip and foot positions precisely with a Vicon motion tracking system, allowing us to study the effect of speed on selection performance in XR in depth. The right-hand side shows the distribution of the final selection points on the targets by index of difficulty. With a higher running speed, participants' selections tend to have a greater dispersion, treating the targets as larger than they actually are.

are not necessarily on par with reference tracking systems, which means that the results might be confounded by tracking issues.

Addressing these gaps, we present a controlled study of direct hand selection in XR, guided by ISO-9241:411 [19], exploring the effect of movement across movement scenarios such as standing still and running (4.5 and 9 km/h). We use a high-precision external motion capture system from Vicon to mitigate the limitations of headset tracking, establishing a 'ground truth' for future XR interaction research on interaction during movement. We not only report on canonical Fitts' law measures such as *MT* and *ER* for comparability [12], [19], but also conduct further analysis on selection movement, cadence, and foot strike to explain the observed performance degradation.

Our results show that higher running speeds negatively impact selection time and throughput, and substantially increase errors, with the error rate exceeding 30% at 9 km/h. Endpoint dispersion along the task axis also grows correspondingly, which drives the effective index of difficulty downward despite greater perceived difficulty. Rather than affecting the initial reaching movement to the target, the added time cost is most visible in the attempts to refine and correct the selection (roughly doubling at 9 km/h). Cadence interacts with these effects: at walking speed, higher cadence reduces errors, but the benefit diminishes at higher running speeds. We summarise our contributions as follows:

- We present results for the first Fitts' law study at true running speeds with precise motion tracking, which helps understand the effect of movement and establishes a 'ground truth', based on which we argue that direct touching is inappropriate for running.
- We provide a better understanding of the reason behind the observed performance drop via further analysis of selection movement, cadence, and foot strike status. Our findings suggest that the source of instability is due to the increased oscillation at higher running speeds, overpowering natural motor control.

Together, these results extend Fitts' law to more intense

movement and provide further insights for designing interaction with XR headsets in motion.

II. RELATED WORK

A. XR Interaction in Motion

The advancement of immersive technologies has brought rich affordances for people to interact with a virtual environment [20], [21]. Examples include using controllers [11], hands [22], feet [23], the head [24], or gaze [25], [26] for target acquisition and selection. While all of these interactions require the user to partially move their body parts, the main body remains relatively static during interaction [4]. Prior research on interaction during movement has investigated XR for in-situ use in scenarios such as sports and exertion, ranging from cycling [26], [27], tennis [28], skiing [29], [30], basketball training [31], indoor workouts [32], and fast walking [1], [33]–[35]. While users benefited from the in-situ feedback, much of this research body did not focus on interaction techniques.

Prior research investigated interaction scenarios where users are in motion. For example, walk- or jog-in-place methods in VR have been studied as input to control navigation and steering [33], [36], [37]. Müller et al. [38] supported target selection by having the user walk on the corresponding projected lanes, which represented different selection options. Sedighi et al. [39] compared the effect of using smart glasses, a smartphone, and a paper notebook on gait variability. Participants were asked to perform different cognitive tasks while walking on a treadmill. Results showed that smart glasses had more negative impacts on stability and gait variability. Lu et al. [24] compared the effect of head- and eye-based gaze selection against eye-based dwell selection on participants' capability to keep pace with a virtual avatar at a walking speed while performing discretionary and monitoring tasks. While the interaction techniques did not differ in walking performance, head- and eye-glance supported faster information acquisition. In another study, voice-based interaction was proposed for drafting emails while walking with a pair of AR glasses [40]. Lu et al. [15] observed that walking speed has a significant

influence on ray-casting selection performance in VR. As users walked faster on a treadmill, their selection errors and time increased, while the ease of selecting dropped. Selection times also changed with speed; higher speeds tended to prolong the time taken for precise target selections. The authors also found that smaller and farther targets were disproportionately harder to acquire while moving, reflecting reduced stability and precision under motion [15]. Li et al. [16] conducted a study in AR, varying user motion from standing to walking, fast walking at a speed of 6 km/h, and jumping, and found a constant decline in selection performance as motion intensity increased. In their experiments, users selecting virtual targets while fast walking or jumping had significantly longer selection times and higher error rates than when stationary or walking at a mild pace, for both direct touch and raycasting selection techniques. In Li et al.'s fast walking and jumping conditions [16], participants' degraded performance was accompanied by reports of greater physical effort and difficulty in maintaining accuracy.

While a large body of research focused on interaction in walking, only a few explored interaction for more intensive locomotion scenarios, such as running. Gallo et al. [41] compared head-gaze interaction, coupled with haptic and audio feedback, against voice interaction, for navigational queries during running activities, revealing user preferences over head-gaze with haptic feedback. Li et al. [35] studied positioning strategies of immersive visualisations, such as head- and body-locked ones, as well as world-anchored embedded visualisations for interval running, altering speeds between 9 and 11 km/h. Results showed that body-locked and world-anchored embedded visualisation achieved better performance while reducing cognitive loads. Similarly, in another study, viewpoint-independent objects were found to be easier for selection while walking fast and jumping [16]. Our study builds on some of the insights from the mentioned previous research to position virtual targets in a world-anchored manner, i.e., independent of the movement of the user, to achieve better selection performance and lower cognitive costs.

B. Fitts' Law in XR

Fitts' Law is foundational in HCI for designing efficient interaction techniques, predicting target selection time [12]. It is commonly applied in the 'Shannon formulation' [13]:

$$\text{MovementTime (MT)} = a + b \times \log_2(A/W + 1),$$

where a and b are constants derived empirically via linear regression, and A and W refer respectively to the movement amplitude and the target size. The index of difficulty (ID) is expressed through the logarithmic term $\log_2(A/W + 1)$, indicating the task difficulty. Based on Fitts' Law, the International Standard ISO 9241-411 [19] was designed to guide the evaluation of the interaction efficiency of new interactive systems. In a nutshell, the ISO definition of a Fitts' study requires participants to select sequentially numbered spherical targets across a circle, representing UI elements. The targets are equally spaced so that the movement amplitudes are equal to the diameter of the circle. The target for selection is highlighted, and a session starts after the participant points to the topmost

target and ends when the sequence is completed (back) at the topmost target. The test is conducted in repetitions with a range of difficulties, varying, e.g., the size of and the distance between the targets. The ISO standard [19] also introduces the use of effective measures: effective movement amplitude (A_e) and effective target width (W_e). W_e is derived from the standard deviation of the distance between the target centre and the selection position (SD_x), with a multiplicative factor of 4.133. Together, A_e and W_e adjust the task's difficulty based on the user's demonstrated skill, reflecting the user's true performance. The adjusted difficulty is referred to as the effective index of difficulty (ID_e), which can be computed via:

$$ID_e = \log_2(A_e/W_e + 1).$$

Finally, the standard defines throughput (TP_e) based on the effective measures as:

$$TP_e = ID_e/MT$$

In XR, Fitts' Law has been primarily examined in stationary environments [10], [12], [22], [42]–[46]. Prior research identified discrepancies in user performance between real and virtual environments, particularly highlighting limitations in VR and AR due to factors such as visual fidelity, the absence of tactile feedback, and the vergence-accommodation conflict [10], [22], [42], [43], [45]. Correspondingly, ample research has proposed methods to improve selection efficiency, such as expanding the size of the target upon movement towards it [44], [47], and filtering on the spatial referencing frame [48]. A recent systematic review of 119 publications synthesises common practices and variability in 3D XR Fitts' studies [12]. The synthesised studies had a median throughput around 3.25 bits/s for HMDs, MTs clustering near 0.5–2.0s, and ER around 0–20%. Most studies chose IDs between 2 and 4.5 bits, leaving higher task difficulties relatively unexplored. Moreover, virtual-hand (36%) and ray-casting (28.6%) techniques were most commonly studied. That survey paper also points out methodological inconsistencies in past studies, where reported units for MT and precision of ID/TP are often non-uniform, complicating cross-study comparison and sometimes producing misinterpreted regression slopes. Informed by this, we adopted ISO-recommended effective measures and reported units consistent with the recommendations from the survey [12] to support comparability with prior research.

While most of Fitts' law studies in XR were conducted in static setups, we identified five key studies that are highly relevant to our study, as summarised in Table I. We discuss these studies in detail and distinguish our contribution clearly below: Zhou et al. [14] examined the effects of walking path curvature on selecting out-of-sight targets in Mixed Reality (MR). They reported decreased selection accuracy with walking and increased path curvature. Head-anchored reference negatively impacted the accuracy. Targets were selected with grasping gestures, which could lead to the Heisenberg effect [49]. Lu et al. [15] investigated how treadmill walking speed affects controller ray-casting to world-fixed and body-fixed targets in VR. They found that higher walking speed increased MT and ER , disproportionately affected small and far targets, and that the body-fixed content interacted more

Study	Interaction Modality	Locomotion Pattern	Tracking Method	Environment	ISO-9241 Task	Index of Difficulty
Zhou et al. 2020 TVCG [14]	Eyes-free Direct Hand Tapping	Walking (≈ 2.5 km/h)	External mocap (OptiTrack & VR Gloves)	VR (Oculus Go)	No	N/A
Lu et al. 2022 IEEE VR [15]	Controller Ray-casting	Walking ($\approx 3, 4$ and 5 km/h)	Headset	VR (Meta Quest 2)	No	N/A
Li et al. 2024 IHCS [16]	Direct Hand Tapping, Hand Ray-casting	Static, Walking (3 km/h), Fast Walking (6 km/h), Jumping	Headset	Optical AR/MR (Microsoft HoloLens 2)	Yes	2.58 to 3.73 bits
Li et al. 2025 CHI [17]	Direct Hand Tapping, Hand Ray-casting	Walking (participant-dependant speed)	Headset	Passthrough AR/MR (Meta Quest 3)	Yes	0.46 to 1.17 bits
Shin et al. 2025 IHCS [18]	Hand Ray-casting	Walking (≈ 4.3 km/h)	Headset	Optical AR/MR (Microsoft HoloLens 2)	Yes	2.54 bits
Sünderkamp et al. 2025 VRST [46]	Gaze + Pinch, Hand Ray-casting, Virtual Hand	Walking (≈ 2.74 - 2.88 km/h)	Headset	Passthrough AR/MR (Meta Quest Pro)	Yes	2.85 to 4.16 bits
Our work reported here	Direct Hand Tapping with Controller Confirmation	Static, Walking (4.5 km/h), Running (9 km/h)	External mocap (Vicon Tracking)	Passthrough AR/MR (Meta Quest 3)	Yes	2.25 to 4.21 bits

TABLE I
COMPARISON OF PRIOR FITTS LAW STUDIES WITH OUR STUDY

strongly with motion than world-fixed content. The latter was also confirmed in Li et al.'s study [16], where they evaluated the effect of selection technique and spatial referencing frames on selection efficiency when walking, fast walking, and jumping. Moreover, the virtual hand technique was found to be less time-consuming and more accurate than ray-casting. However, in their study, the walking speed was set to 3 km/h, and although they claimed their fast walking condition at 6 km/h as running, it only qualifies as fast walking, as previous work has identified the minimum transition speed from walking to running to be 7.2 km/h [50], [51]. The superiority of the virtual hand technique was also observed in Shin et al.'s study [18], which explored ray-casting selection at walking speed with a Fitts' tapping task. They compared ray-casting and clicking/dwelling with eye gaze, hand, and head, respectively. Gaze was found to be more efficient during mobile scenarios, while clicking was more efficient than dwelling. Li et al. [17] quantified situational impairments by studying walking with an encumbrance in MR while performing direct selection, ray-casting, and text entry. They reported substantial performance degradation during walking and additional losses due to the encumbrance, with interaction effects amplifying the costs. Sünderkamp et al. [46] studied various distant selection methods using a Fitts' tapping task while participants walked outdoors. Their findings suggested gaze with pinch selections worked better than raycasting and the virtual hand techniques.

In comparison to these key prior works, which studied the effect of locomotion up to walking speed, we extended our investigation to include a true running speed. Motivated by results from Li et al.'s [16] and Shin et al.'s study [18], we studied only direct hand tapping as it had been proven superior in their context. In the direct hand tapping condition in Li et al.'s studies [16], [17], a selection was made only when the selecting finger touched a target [17] or the selection plane [16], which may result in unintentional selections and skew the recorded *MT* and *ER*. To avoid this, we asked our participants

to press a button on the controller to confirm a selection. In our study, a controller was held in the participants' non-pointing hand to avoid the Heisenberg effect [15], [49]. Moreover, no prior work explored the potential reasons that could explain the degraded performance further, something that we aimed to investigate through the use of a precise external tracking system.

C. Pointing Movement Modelling

In 1899, Woodworth [52] modelled a rapid aimed movement with two phases: ballistic and correction. The ballistic phase corresponds to the initial impulse that propels a movement towards a target, whereas the correction phase refers to the perceptually controlled fine-tuning stage that corrects any unintentional errors. There are two approaches to separate the two submovements. The first approach was to identify the moment in the movement trajectory where the target surface was tangent to the selector as the boundary between ballistic and correction phases [53]. However, this approach does not reflect human motor control well [54]. Alternatively, the velocity of the movement can be used as a parsing criterion for movements. Meyer et al. [55] proposed detailed parsing criteria for the two distinct submovements. First, the movement trajectories were smoothed and differentiated as a function of time. Then, the end of the primary submovement was labelled as either a type 1, type 2, or type 3 submovement. Type 1 submovement refers to the returning movement to the target after overshooting, which occurs when the velocity crosses zero from positive to negative [55]. Type 2 submovement refers to scenarios when an undershoot to the target happens, in which the acceleration crosses zero from negative to positive [55]. Type 3 submovement refers to the situation when the movement is slowing down towards the target, in which a zero jerk is reached from positive to negative [55]. However, these criteria were initially proposed for one-dimensional movements and require adjustments to suit three-dimensional

pointing [55], [56]. Nieuwenhuizen et al. [57] proposed an extended set of parsing rules that separated a 3D pointing movement into five phases. They introduced the latency and verification phases to address the insignificant movements that happen at the start and the end of a movement. After the latency phase, the initiation phase was proposed before the ballistic and correction phases would happen, to address the start of the actual movement [57]. Liu et al. [56] proposed a modification to Meyer's original parsing criterion [55]. They further studied the difference in target selection time between the virtual and physical environments. Their results showed that selection time in virtual environments was twice as slow as in the physical environment – the duration of ballistic movement in the virtual environment was nearly double that of the physical, whereas the correction phase was six times longer in the virtual environment than in the physical world [56]. A more recent study by Deng et al. [54] inherited Liu et al.'s modification [56] with a more generous criterion for the type 3 movement, which is identified once the acceleration was negative and its absolute value became less than one tenth of its maximum. They also further divided ballistic movement into acceleration and deceleration phases [54]. In our research, we took Deng's parsing criteria to segment movements, but collapsed acceleration and deceleration into the ballistic phase.

III. MOTIVATION AND HYPOTHESES

XR has been applied to various motion-intensive data visualisation scenarios. Still, these applications often focused on the visual feedback, lacking support for interaction, which could allow people to gain further insights about the data. While pointing and selecting interactions have been well-studied in XR with Fitts' Law studies [12], only a few studies have evaluated their efficiency in mobile scenarios [14]–[18], [46]. Prior work reported that mobility harmed interaction, especially with ray-based techniques [16] and head-anchored content [14]–[16], [46]. However, not all prior studies [14], [15] followed the Fitts' tapping task methodology standardised in ISO-9241 [19]. The studied movement intensity was up to a fast walking speed (6 km/h in Li et al.'s research [16]). Moreover, findings were primarily on the effect of speed on selection performance, without further exploration of the potential causes. Conversely, our study advances the literature by (i) testing selection during a *real running speed* [50], [51] while following the ISO-9241 Fitts' tapping methodology [19], and (ii) through tracking with an accurate external mocap system, revealing some of the causes that explain the degradation of selection performance, through movement separation and cadence analysis, offering a stronger foundation for designing motion-aware XR selection interaction. We thus hypothesise:

- **H1. Movement speed affects MT; the higher the speed/cadence, the longer it takes to reach a target** — While partially proven by prior works [14]–[18], [46], we seek to extend this to true running speeds.
- **H2. During fast movement, it is the correction phase at the end of a selection, rather than the ballistic movement of the arm, that causes the effect on MT** — No prior studies about Fitts' tapping tasks in moving scenarios investigated selection movements further.

Nonetheless, a prior study with a static setup evidenced that the correction phase increases drastically in virtual environments relative to physical counterparts [56]. We thus assume that a similar observation would be obtained when movement intensity increases.

IV. USER STUDY

We conducted a study to (1) evaluate the effect of intense movement on the typical Fitts' Law equation (**H1**), and (2) to understand the potential reasons behind the observation (**H2**).

Participants — At the recruiting stage, participants filled out a demographic questionnaire. Participants were recruited only if they had experience with running. We ended up recruiting 22 participants, including eight females and fourteen males. Thirteen of the participants were aged 25-34 years, five 35-44 years, three 18-24 years, and one was 45-54 years. All participants were frequent runners who ran weekly to varying degrees – ten ran up to one hour per week, five 1-3 hours, four 3-5 hours, and three ran 5-7 hours. All participants were right-handed and had normal or corrected-to-normal vision. Participants also indicated their experiences with running and with immersive technologies such as XR headsets using a 5-point Likert scale (ranging from 1 to 5, no experience to expert proficiency), resulting in an average score of 3.05 ± 0.71 for running, and 3.09 ± 1.35 for XR.

Apparatus — We implemented a Fitts' Law tapping task following ISO 9241-411 [19] in Unity version 6000.0.37f1³. The implementation was deployed on a Meta Quest 3 Mixed Reality headset⁴ in passthrough mode. We set up a Vicon Motion Tracking system⁵, composed of six motion cameras, tracking at 220 Hz the positions of the fingertip of the right index finger and the feet (see Figure 1). We 3D-printed three rings, fitting small, medium, and large circumferences of the index finger, as well as two shoe clips to mount the Vicon trackers (see Figure 1). We uniquely attached four Vicon tracking balls to each print, forming rigid body positions for the tip of the index finger and the feet in Vicon. The positions were transferred to the headset in real-time through a UDP protocol. Since Vicon and the headset use different coordinate systems, a Procrustes alignment [58] was used to align the positions tracked by the Vicon system with the coordinate system of the headset.

Procedure — As depicted in Figure 3, the researcher first performed a calibration to ensure the fingertip and feet positions were properly aligned. Next, participants were informed of the research's purpose and procedure, and they signed a consent form. Afterwards, participants warmed up on the treadmill at a speed of their choice for two minutes. The speed of the treadmill was closely monitored by the researcher. After that, the researcher helped them don the headset, ring, the shoe clippers, and adjust the interpupillary distance. Participants then recalled their viewing angle (how far their head was tilted down) while running on the treadmill. The target plane appeared at a fixed position slightly in front of the treadmill

³Unity: <https://unity.com/releases/editor/whats-new/6000.0.37f1>

⁴Meta Quest 3: <https://www.meta.com/au/quest/quest-3>

⁵Vicon: <https://www.vicon.com>

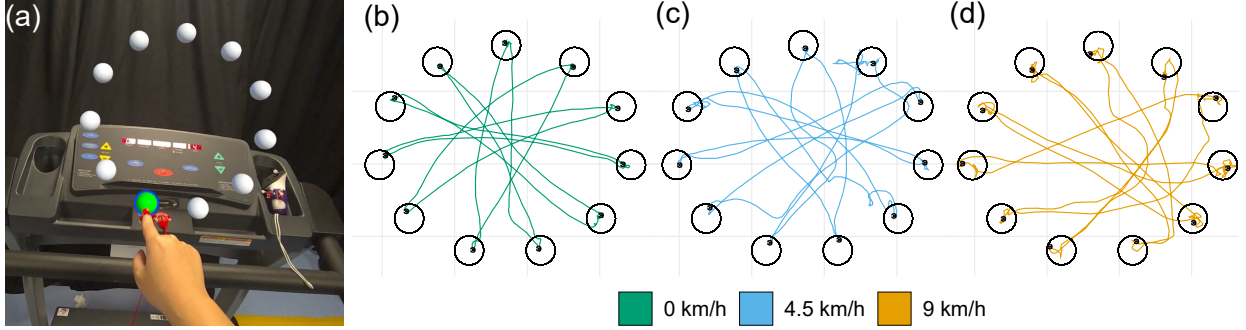


Fig. 2. (a) When the fingertip reaches a target, an outline appears around it. (b-d) Example movement trajectory along X and Y axes at the studied speeds (taken from P5, ID ≈ 3.29). Solid dots represent the position where a selection was made.

panel, with the height adjusted to the participant's chest level (25 cm below the head position), facing the participant perpendicularly, based on the recorded viewing angle. After that, they spent around 2 minutes getting used to the visuals, both standing still and running at their chosen speed, while freely practising the tapping task. Participants used the tip of their right index finger to point at a target while they held a controller in their left hand, with the trigger button pressed to confirm a selection, thereby mitigating the Heisenberg effect [49], [59]. Controller confirmation also isolated the failures that can occur with spatial pointing with gesture recognition, as optical pinch or hand-gesture recognition can degrade when the hand moves relative to the headset during running. This choice prioritised internal validity and should be interpreted as a controlled benchmark rather than a deployable technique for real-world use.

Eleven spheres were presented in a circle according to the multidirectional selection task 4.3 in ISO 9241-411 [19]. Participants were told to make sequential selections of each target currently highlighted in green as quickly and accurately as possible. When the participant's fingertip collided with a target, it was outlined in blue to provide visual feedback (see Figure 2a). A selection was then deemed to be successful if the trigger button of the controller was pressed while the fingertip (still) collided with the target. When an unsuccessful selection happened, a short beep sound was played to indicate the error. A sequence always began with the participant selecting the topmost target and continued with the next diagonal target in clockwise order, until it finished at the topmost target. Participants performed the same selection tasks while standing still, and at speeds of 4.5 km/h and 9 km/h, respectively. After each speed, participants completed a questionnaire, consisting of a NASA-TLX [60] to assess their cognitive workload, and a Borg CR-10 scale [61] to record their level of exertion. Then, participants were asked to rest fully until their exertion level

had become minimal, after which they returned to the treadmill for the next condition.

Study design — We designed a 3×9 within-subject study, varying the running speed and the index of difficulty of the selection task. The running speed included a condition where the participant stood still, which served as the baseline; a lower (fast walking) speed of 4.5 km/h; and a higher (running) speed of 9 km/h. We chose three target sizes: 2 cm, 3 cm, and 4 cm, and three distances between the targets: 15 cm, 25 cm, and 35 cm. These formed 9 unique indices of difficulty ranging from 2.25 to 4.21, covering common difficulties reported in prior research [12]. Each participant completed 27 trials, each with a unique combination of the 2 independent variables. The order of the trials was balanced using a Latin Square to mitigate learning effects. The study took around 45 minutes for each participant.

For objective data, we collected the angle at which the targets were tilted down, which was consistent per participant across all speeds; per-selection *MT* (seconds); *ER* (%); selection *offset* (meter); *SD_x* (meter); *ID_e* (bits); *TP_e* (bits per second); *cadence* (steps per minute); and per-frame movement data for the fingertip and feet. *MT*, *ER*, and *cadence* are directly collected measures, whereas selection *offset*, *SD_x*, *ID_e*, and *TP_e* are derived and can reflect the endpoint variability and speed-accuracy trade-offs. Note that the way we measured selection *offset* and *SD_x* was different. Selection *offset* was calculated based on the distance between the fingertip surface and the target surface, whereas *SD_x* was measured based on the distance between the center of the fingertip and the target's center. For subjective data, we collected cognitive task load using the NASA-TLX [60] and exertion level using a Borg CR-10 scale [61], both of which were self-reported by participants.

V. RESULTS

We collected $3 \times 9 \times 11 \times 22 = 6534$ trials of selection, as well as the per-frame movement of the selecting index finger

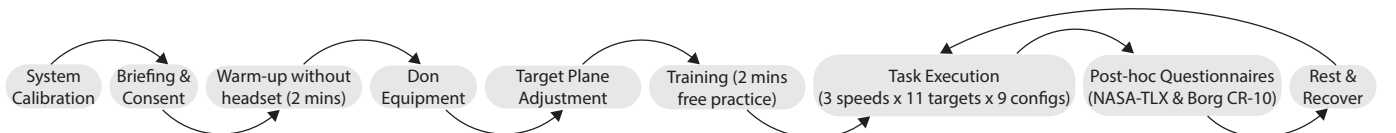


Fig. 3. Study Procedure

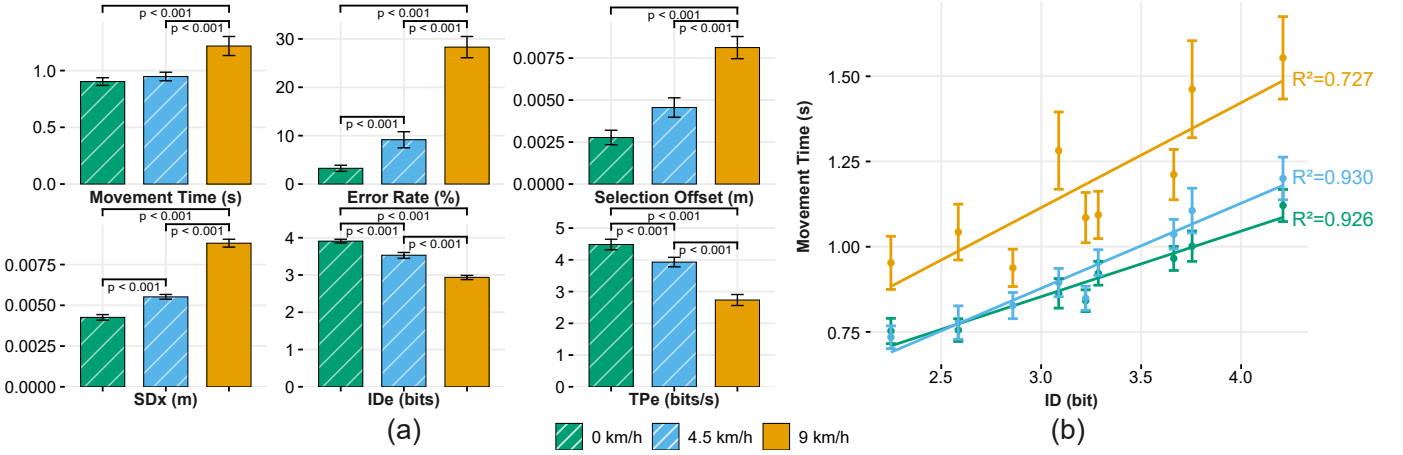


Fig. 4. (a) Effect of running speed on *MT*, *ER*, selection *offset*, *SD_x*, *ID_e*, and *TP_e*. Striped bars represent speeds similar to prior studies. (b) Effects of speed on *MT* over *ID*. Note that the unmatched conditions with 4 cm of size and 25 cm of distance (*ID* = 2.86) at 4.5 and 9 km/h were included in the figure but excluded from all other analyses.

and both feet at 72Hz. Due to an implementation issue, the system presented the condition of 4 cm size and 15 cm distance twice at 0 km/h, but the intended condition with 4 cm size and 25 cm distance at that speed was not administered. To retain a fully crossed design for inferential analyses, we (i) removed the duplicated condition at 0 km/h ($22 \times 11 = 242$ trials) and (ii) omitted the unmatched conditions with 4 cm of size and 25 cm of distance at 4.5 and 9 km/h ($2 \times 22 \times 11 = 484$ trials). We excluded two trials from P6 due to brief initialisation transients in the tracking signal, as well as 27 trials that were considered unintentional selections or double clicks. The final analysis set comprised 5779 trials. For completeness, unaffected conditions at 4.5 and 9 km/h are shown in the figures, and the exclusions were applied uniformly to maintain comparability across speeds.

We analysed the data with R using Repeated Measures (RM) ANOVA. We ran a Shapiro test [62] to verify the normality of the collected data, including *MT*, *ER*, *Offset*, *SD_x*, *ID_e*, *TP_e*, *NASA-TLX*, and *Borg CR-10* scores. Neither the original nor the log-transformed forms of these data were normally distributed. For the objective data, we thus applied rank-based Inverse Normal Transformations (INT) [63] to the original data before conducting an ANOVA. Before these ANOVAs, we assessed the assumption of sphericity using Mauchly's test, confirming that the assumption was met for all corresponding metrics. We then conducted post-hoc analyses with 95% confidence intervals on all significant main effects revealed by ANOVA and adjusted the results using the Bonferroni correction. We report the partial η^2 (ETA squared) in ANOVA as the indicator for the effect size. For the subjective data measured by *NASA-TLX* and the *Borg CR-10* scale, we ran a Friedman test, followed by a Nemenyi test for post-hoc analysis to determine if there is a significant main effect. We reported Kendall's *W* to indicate the effect size. We then conducted a Fitts' law analysis for the independent variable of speed, followed by a detailed analysis that further explores the mechanism behind the observations.

A. Statistical Analysis

Directly collected measures:

MT — Speed had a significant effect ($p < 0.001$, $F = 27.48$, $\eta^2 = 0.57$) on participants' movement time to finish a selection. Post-hoc analysis further revealed a significant difference between 0 and 9 km/h ($p < 0.001$), and between 4.5 and 9 km/h ($p < 0.001$).

ER — Speed had a significant effect ($p < 0.001$, $F = 108.26$, $\eta^2 = 0.84$) on selection error rate. Significant effects were found between all pairs of speeds.

Offset — We kept only false selections before performing INT and ANOVA. Results showed that speed had a significant effect ($p < 0.001$, $F = 30.654$, $\eta^2 = 0.59$) on the offset between participants' fingertip position and that of the selection target. Significant effects were observed between the speeds of 0 and 9 km/h ($p < 0.001$), and 4.5 and 9 km/h ($p < 0.001$).

SD_x — Speed had a significant effect ($p < 0.001$, $F = 153.97$, $\eta^2 = 0.88$) on *SD_x*. Significant effects existed between all pairs of speeds.

Tilting Angle — Participants were allowed to freely adjust the tilting angle of the targets before the study. The angle was consistent throughout the study for each participant. Overall, participants averaged a tilting angle of $31.5(\pm 5.7)$ degrees.

Derived measures:

ID_e — Speed had a significant effect ($p < 0.001$, $F = 161.12$, $\eta^2 = 0.88$) on the effective index of difficulty. Significant effects were found between all pairs of speeds.

TP_e — Speed had a significant effect ($p < 0.001$, $F = 87.09$, $\eta^2 = 0.81$) on participants' *TP_e*. Post-hoc analysis showed significant effects between all pairs of speeds.

Subjective measures:

NASA-TLX — Speed had a significant effect ($p < 0.001$, $\chi^2 = 40.364$, Kendall's *W* = 0.92) on the overall task load when performing the selection task. Significant effects existed between all pairs of speeds. We further performed a breakdown analysis for each dimension – as shown in Figure 5a, speed significantly affected participants' 1) mental demand in completing the selection task ($p < 0.001$, $\chi^2 =$

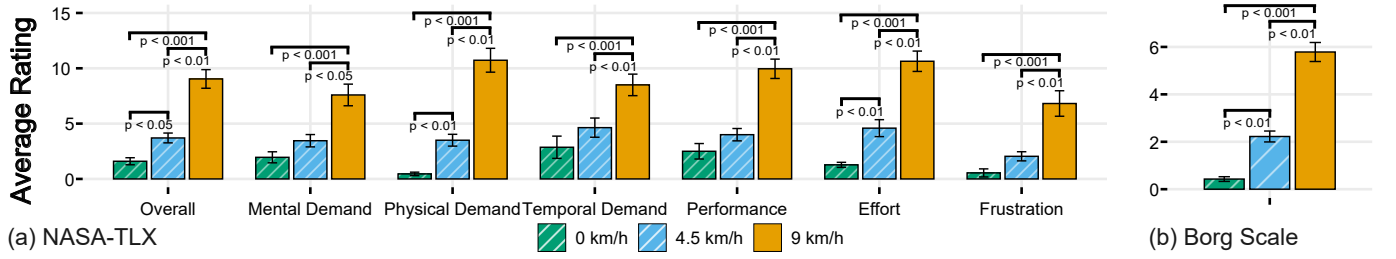


Fig. 5. Average rating of participants' (a) cognitive workload, measured with NASA-TLX, and (b) perceived exertion level, measured with the Borg CR-10 scale [61]. Striped bars represent similar speeds in prior studies.

21.641, Kendall's $W = 0.49$), between speeds of 0 and 9 km/h ($p < 0.001$) as well as 4.5 and 9 km/h ($p < 0.05$); 2) physical demand ($p < 0.001$, $\chi^2 = 41.518$, Kendall's $W = 0.94$), between all pairs of speeds; 3) temporal demand ($p < 0.001$, $\chi^2 = 32.667$, Kendall's $W = 0.74$), between speeds of 0 and 9 km/h ($p < 0.001$), as well as 4.5 and 9 km/h ($p < 0.01$); 4) performance ($p < 0.001$, $\chi^2 = 30.296$, Kendall's $W = 0.69$), between speeds of 0 and 9 km/h ($p < 0.001$), as well as 4.5 and 9 km/h ($p < 0.01$); 5) effort ($p < 0.001$, $\chi^2 = 43.517$, Kendall's $W = 0.99$), between all pairs of speeds; and 6) perceived frustration ($p < 0.001$, $\chi^2 = 34.203$, Kendall's $W = 0.78$), between speeds of 0 and 9 km/h ($p < 0.001$), as well as 4.5 and 9 km/h ($p < 0.01$).

Borg CR-10 Scale — Speed had a significant effect ($p < 0.001$, $\chi^2 = 44$, Kendall's $W = 1$, Figure 5b) on participants' perceived level of exertion for the selection task. This existed between all pairs of speeds.

B. Fitts' Law Analysis

We performed a Fitts' law analysis on the speeds in our study. As shown in Table II, we found strong fits for standing still and walking ($R^2 = 0.926$ and 0.930), while a weaker fit was found for the running speed at 9 km/h ($R^2 = 0.727$).

To further examine the lower fits at 9 km/h (Figure 4b), we fitted a Linear Mixed-Effects Model (LMM) for MT , with Speed, ID , and their interaction as fixed effects, and participant as a random intercept. The analysis revealed a significant $Speed \times ID$ interaction ($\chi^2(2) = 6.58$, $p = 0.037$). A post-hoc analysis with Bonferroni correction confirmed that the slope of MT relative to ID became significantly steeper at 9 km/h compared to standing at 0 km/h ($p = 0.032$).

C. Explorative Analysis

To further understand the potential reasons behind our observations, we conducted analyses on factors that might have

Speed (km/h)	Coefficients	Fit (R^2)	AIC Score
0	$MT = 0.276 + 0.192 \times ID$	0.926	-27.340
4.5	$MT = 0.138 + 0.248 \times ID$	0.930	-23.821
9	$MT = 0.269 + 0.289 \times ID$	0.727	-8.546
Overall	$MT = 0.228 + 0.243 \times ID$	0.454	-15.806
Overall (w/ Speed)	$MT = 0.061 + 0.247 \times ID + 0.035 \times Speed$	0.821	-41.727

TABLE II
FITTS' LAW ANALYSIS RESULTS.

an impact on selection performance: 1) fingertip movement phases, 2) cadence, and 3) the correlation between foot strikes and selection timing.

Fingertip Movement Phases — We followed the three-phase segmentation for selection movements in immersive environments, proposed by Deng et al. [54], to handle the recorded per-frame fingertip movements. The model decomposed a selection movement into acceleration, deceleration, and correction. We combined the acceleration and deceleration phases to form the ballistic phase. We computed the velocities of the fingertip along three axes and applied a 4th-order Butterworth low-pass filter with a cutoff frequency of 9 Hz to mitigate noise introduced by running and hand jitters [54]. To separate the phases, we marked the closest time frame when the total selection speed became less than half of the average maximum selection speed, and met one of the three conditions: 1) the selection speed changed to negative from positive, 2) the acceleration changed to positive from negative, and 3) the acceleration was negative with its magnitude being less than 10% of its maximum. Movements before that time frame were labelled as the ballistic phase, while those after were labelled as the correction phase (see Figure 6 for instance).

Because the data were not normal, we INT-transformed the ballistic and the correction times and ran RM ANOVA on each. Before ANOVAs, we ran Mauchly's tests and confirmed that the assumption of sphericity was met by the data for both the ballistic and the correction phase. Results showed that speed had no significant effect on the ballistic phase, whereas a significant effect ($p < 0.001$, $F = 50.285$, $\eta^2 = 0.71$) was found on the correction phase. A post-hoc pairwise comparison revealed that a significant effect existed between all pairs of speeds. As shown in Figure 7bd, the effect on the correction

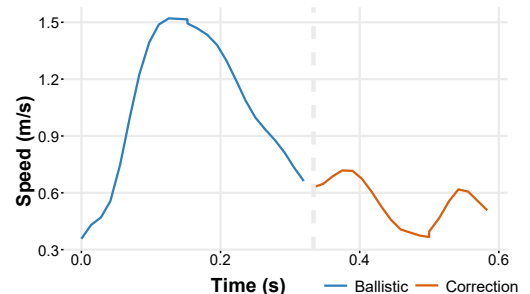


Fig. 6. Ballistic/correction phase segmentation of an example selection (taken from P7, Speed = 9, $ID \approx 4.21$).

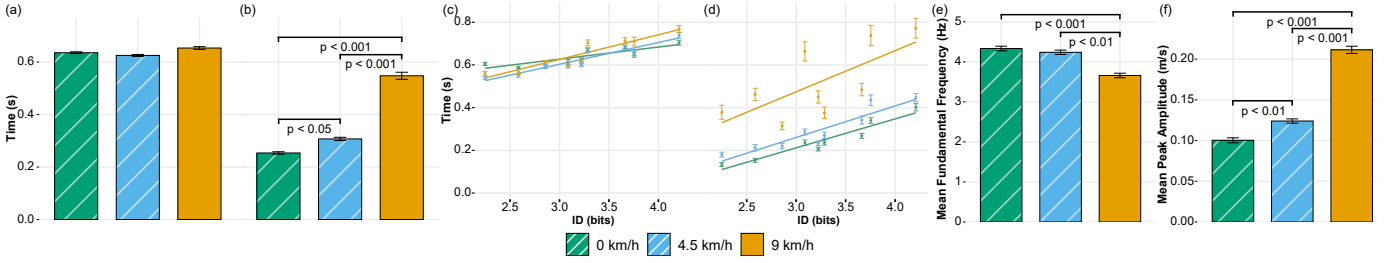


Fig. 7. Effects of speed on (a) the ballistic phase, and (b) the correction phase, as well as over ID on (c) the ballistic phase, and (d) the correction phase. Note that the unmatched conditions with 4 cm of size and 25 cm of distance ($ID = 2.86$) at 4.5 and 9 km/h were included in (c) and (d) but excluded from all other analyses. (e) Mean fundamental frequency, and (f) mean amplitude of the correction phase. Striped bars represent similar speeds in prior studies.

phase at the speed of 9 km/h was particularly strong, doubling the time needed for correction over 0 and 4.5 km/h.

We conducted a frequency analysis on the correction movement. We computed the mean fundamental frequency of the selections with a fast Fourier transform [64], and the mean movement amplitude, based on the height of the local maxima. We INT-transformed both the fundamental frequency and the amplitude before applying RM ANOVA for the speed. We found that speed had a significant effect on both the movement frequency ($p < 0.001$, $F = 12.136$, $\eta^2 = 0.37$) and amplitude ($p < 0.001$, $F = 41.145$, $\eta^2 = 0.66$) in the correction phase. As shown in Figure 7a-d, post-hoc analyses revealed that significant differences in the fundamental frequency of speed existed between all pairs of speeds except between 0 and 4.5 km/h (Figure 7e), and significant differences in movement amplitude existed between all pairs of speeds (Figure 7f).

Cadence Analysis — Through visual inference (see Figure 8a), we identified two distinct movement patterns at the speed of 4.5 km/h. While the majority of participants had a typical walking cadence, four of them chose to run and thus ended up with a cadence higher than 140, which is the transitioning cadence from walking to running as reported in prior work [50], [51]. We excluded these four data points as outliers. To investigate the effect of cadence on selection performance during movement, we employed Generalised Linear Mixed Models (GLMMs) to account for missing data points and the non-normal distribution of the dependent variables. We then modelled each dependent variable as a function of speed (4.5 km/h and 9 km/h), relative cadence (binned as 'high' or 'low' within each speed), and their interaction. Our analysis revealed a significant interaction effect between speed and cadence on ER ($\chi^2 = 8.364$, $p < 0.01$). Post-hoc comparisons showed

that when running with 4.5 km/h, the higher cadence resulted in a significantly lower ER compared to lower cadences ($z = 3.317$, $p < 0.01$). No significant difference was found in ER between high and low cadences at 9 km/h. We also found a significant interaction effect between speed and cadence on selection *offset* ($\chi^2 = 6.630$, $p < 0.05$). However, post-hoc analysis did not show significant differences between relative cadences within each speed. No significant main effects or interactions involving cadence were found for other performance metrics, such as MT , TP_e , ID_e , or SD_x (all $p > 0.05$).

Foot Strike and Selection — We derived per-frame foot-contact states from the vertical foot trajectories, filtered with a 4th-order Butterworth low-pass filter at a cutoff frequency of 9 Hz to attenuate jitter. We computed the median absolute deviation (MAD) of the troughs as the threshold ground level. This was computed for each foot to account for the potential difference in the placement of the foot trackers. A frame was labelled 'on the ground' if either the height of the left or the right foot was below the MAD threshold. To assign a contact state to each target selection, we matched the selection timestamp to the nearest per-frame label. Out of the 3847 selections during movement (11 targets $\times$ 2 speeds $\times$ 8 IDs $\times$ 22 participants, and disregarding 25 trials that were made out of mistakes) made while in movement, 1371 were made in the air and 2476 on the ground. Again, we used GLMMs to investigate the effect of foot status on selection performance. We did not find any difference in the dependent variables between in-air and on-ground selections (all $p > 0.05$).

VI. DISCUSSION

We first revisit the hypotheses and then discuss our results in comparison to closely related prior works where possible.

H1. Movement speed affects MT ; the higher the speed/cadence, the higher the time it takes to reach a target

— Our results partially supported **H1**. MT increased with speed, and the fit to Fitts' law [13] weakened at 9 km/h (see Table II), while a model including speed as a factor markedly improved overall fit. Empirically, movement speed exerted a strong main effect on selection performance. This is especially evident at higher speeds. Compared to standing still, at a speed of 9 km/h, MT increased 25%; ER surged to more than 30%, more than 6 times higher than when standing still; offset and endpoint dispersion along the task axis (SD_x) doubled; and throughput decreased nearly 50%. This is accompanied by a

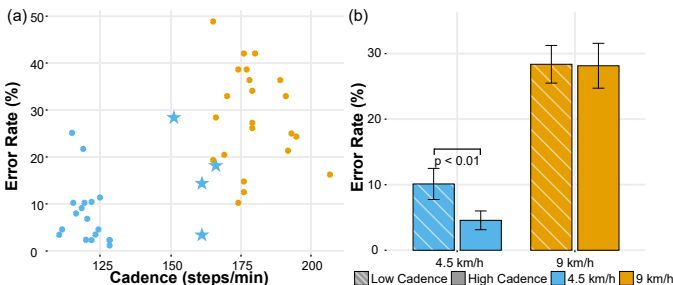


Fig. 8. Effects of cadence on ER . Data points marked as stars in (a) were excluded at 4.5 km/h as outliers.

significant increase in the perceived exertion and task load, as indicated by the Borg scale and NASA-TLX outcomes. Beyond standard changes in target size/distance, the magnitude of these changes collectively indicates that movement speed is a dominant driver of selection performance in motion. However, our results do not suggest the same for cadence. At the same speed, running with a higher cadence interestingly does not prolong the *MT*.

H2. During fast movement, it is the correction phase at the end of a selection, rather than the ballistic movement of the arm, that causes the effect on MT — Our analysis supports **H2**, as the ballistic movement time was statistically stable across speeds, whereas the correction movement time increased significantly, nearly doubling at 9 km/h over 0 or 4.5 km/h. This indicates that running induces more and longer corrective adjustments before a selection is confirmed. Moreover, most selections occurred when a foot was on the ground rather than in the air, suggesting participants tend to exploit gait-induced stability windows to finalise selections. However, this does not help them make more accurate selections.

Our analysis in subsection V-C explains why performance decreases when moving. As speed goes up, and while the time spent reaching a target remains consistent, participants spent a significantly larger amount of time correcting the selection of a target (see Figure 7a-d). Meanwhile, the source of selection errors changes fundamentally with speed. In static and walking speed conditions, selection endpoints were tightly clustered (Figure 1), movements were smooth (Figure 2bc), the fundamental frequency of the correction phase was relatively high (≈ 4.2 Hz), and the amplitude relatively low (≈ 0.1 m/s, Figure 7ef). This is consistent with the behaviour of rapid visuomotor sub-movements [65], indicating that error was primarily driven by motor noise that participants can actively correct. However, at 9 km/h, we observed a distinct ‘gait-dominance’ effect. The fundamental frequency of the correction movement dropped significantly to ≈ 3.5 Hz (Figure 7e), while the amplitude nearly doubled over the other two speeds (Figure 7f). This 3.5 Hz frequency also aligns with the average running cadence (≈ 175 -190 steps/min, 2.9-3.2 steps/s), likely indicating that gait-induced oscillation overpowered natural motor control. This is also evidenced in the greater dispersion of selection endpoints (Figure 1) and the trajectories (Figure 2d), and in our findings about the cadence analysis, which failed to mitigate errors at high speeds (Figure 8b). Further, this physical instability may be compounded by a shift in perceptual difficulty. As shown in the appendix, the estimated user-to-target distance and angle reveal that as speed increases to 9 km/h, the user drifts further backward, effectively shrinking the target visual angles. This reduction in visual angle may also contribute to the higher *ER* and longer correction times observed during the running condition, as participants are forced to select perceptually smaller targets while battling the aforementioned gait-induced oscillations. Together with the increased mental taskload and exertion level (Figure 5), this results in a fatiguing, high-amplitude oscillation and a visually demanding task that exceeds participants’ control. This impact is reflected in the fits of Fitts’ law to speeds. As shown in Figure 4b, while Fitts’ law

remained applicable at 9 km/h, its model fitness dropped from $R^2 = .926$ and $R^2 = .930$ at 0 and 4.5 km/h to $R^2 = .727$ at 9 km/h. This suggests that the prediction of *MT* is no longer dictated by *ID*, but also by the aforementioned factors.

We measured and analysed results with the effective metrics introduced in ISO 9241:411 [19], such as SD_x , ID_e , and TP_e , in which SD_x increases and TP_e decreases as the speed goes up, which is as expected. However, counter-intuitively, the effective index of difficulty, ID_e , falls as the speed increases. The reason behind this is that it is defined as $ID_e = \log_2(A_e/W_e + 1)$. In our study, the effective amplitude (A_e) did not change, as the chosen target distances were consistent across speeds. Meanwhile, the effective target width (W_e), derived from the standard deviation of the distance between the target centre and the selection position, significantly increased, causing the logarithmic term to decrease. The effective width represents the space in which participants actually made their selections. In a typical Fitts’ pointing task, participants are required to balance speed and accuracy. If they prioritise speed, their accuracy will decrease. In our case, in the running conditions, participants increased their selection time but still had to sacrifice accuracy, presumably due to the oscillations incurred while running. As a result, they had to ‘effectively’ treat the target as being larger than intended; thereby the effective index of difficulty of what they are actually selecting decreases. This dissociation reinforces the need to incorporate locomotion terms explicitly into predictive models, rather than assuming ISO effective measures will always track perceived difficulty across different movement speeds. Based on the effective target width (derived from the multiplication of SD_x by 4.133, we make recommendations for the size of the targets: a target size should be at least 0.0176, 0.0228, and 0.0364 meters for it to be efficiently selected at the speeds of 0, 4.5, and 9 km/h, respectively.

We compared our results with prior findings to identify similarities and discrepancies. When static, prior studies [12] suggested that the *MT* reported by existing works concentrates around 1 to 2 seconds, *ER* between 0 and 20 percent, and throughput between 1.5 and 5 bits/s. While we recorded sub-second *MT* (0.9 seconds) and average *ER* as low as 3.36%, our throughput outcomes align with their observation. Similarly, for the walking condition, we observed lower *MT* (averaged at 0.93 seconds) and lower *ER* (averaged at 9.61%) compared to those reported in previous studies with similar conditions: at 3 km/h around 1.1 seconds for *MT*, 18% for *ER*, and 1.25 seconds for *MT*, 50% for *ER* at 6 km/h in Li et al.’s study [16]; at 4.3 km/h, 1.46 seconds for *MT*, 18% for *ER* in Shin et al.’s study [18]. Still, contrasting Shin et al.’s findings [18], our results show no significant difference for *MT* between the speeds of 0 and 4.5 km/h. On the other hand, our *MT*s are closer to those reported in Li et al.’s study [17], where it averaged around 0.8 seconds for direct hand selection while walking. However, in Li et al.’s study, a selection was confirmed with a pinch gesture, while the studied target sizes were much larger (4 cm is the largest size in our study, whereas it is the smallest in theirs), and the distances were much smaller (3 cm and 5 cm compared to 15 cm, 25 cm, and 35 cm in our study). Considering the headset-based tracking

systems used in previous work, we attribute these differences to the substantially higher accuracy and faster tracking rates of the Vicon tracking system that we adopted in our study, which enabled sub-millimeter-level tracking at 220 Hz. However, the findings in our study should be interpreted as the best-case scenario. Translating these findings to commercial optical see-through or passthrough headsets requires careful consideration of the hardware limitations inherent to the built-in tracking, as subtle increases in latency and spatial jitter will compound target selection performance [66].

It should also be noted that our study prioritised experimental control over external validity. First, treadmill running differs from overground running through the belt motion, constrained position, and the absence of terrain changes, which may alter cadence, vertical oscillation, and attentional demands. Our controlled laboratory study also does not reflect the dynamically changing scenario in actual outdoor running, where performing a tapping task may incur safety concerns such as tripping hazards. Second, controller confirmation reduced unintended selections and gesture-recognition confounds, but it is less ecologically valid than bare-hand confirmation for lightweight running interfaces. Third, regarding the Fitts' law analysis performed in subsection V-B, while the unmatched condition ($ID = 2.86$) was appropriately excluded to maintain symmetry for the ANOVA, its absence still does not meaningfully skew the reported linear regression slope, as the trend remains robust, evidenced by the remaining IDs distributed around the trend lines in Figure 4b. Finally, the reported target distance and visual angle are estimates, as we did not record head position. While it captures horizontal treadmill drift, it does not account for the runner's forward trunk lean at higher speeds, which may result in the true visual angles during the 9 km/h condition being marginally larger than estimated.

VII. CONCLUSION

We demonstrate that running speed significantly degrades target selection performance in XR, with error rates exceeding 30% at 9 km/h alongside increased cognitive load and physical exertion. While the ballistic reaching phase remains stable across speeds, the correction phase doubles in duration as participants struggle to fight against the vertical oscillation. While a higher cadence reduces errors during walking, it provides no benefit during running, and timing selections with foot strikes on the ground offers no advantage over mid-air selections. Still, potential solutions such as fatigue-aware interactions [67] and movement filtering [48] could be utilised to improve selection efficiency in future running systems.

APPENDIX

Target Distance and Visual Angle — To address the dynamic nature of the user-to-target distance caused by positional drift on the treadmill, we estimated the target distance and visual angles [68] at the moment of selection. As per-frame head tracking was not logged, we utilised the midpoint between the left and right foot coordinates as a proxy for the head's horizontal (X, Z) location [69]. Combined with

the calibrated, static vertical (Y) offset of 25 cm from the target plane, we computed the distance to the target and the corresponding visual angle, as summarised in Table III.

Speed	Distance	Visual Angle (°) by Target Size		
		2 cm	3 cm	4 cm
0 km/h	50.70 ± 6.76 cm	2.29 ± 0.30	3.47 ± 0.50	4.59 ± 0.67
4.5 km/h	52.23 ± 8.00 cm	2.26 ± 0.37	3.38 ± 0.54	4.44 ± 0.76
9 km/h	63.34 ± 7.40 cm	1.82 ± 0.23	2.78 ± 0.36	3.65 ± 0.44

TABLE III
ESTIMATED MEAN TARGET DISTANCE UPON SELECTION AND VISUAL ANGLES ACROSS SPEEDS AND TARGET SIZES ($M \pm SD$).

REFERENCES

- [1] A. Li, C. Perin, G. Demartini, S. Viller, J. Knibbe, and M. Cordei, "Running with data: a survey of the current research and a design exploration of future immersive visualisations," *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–11, 2025.
- [2] A. Islam, A. Bezerianos, B. Lee, T. Blascheck, and P. Isenberg, "Visualizing information on watch faces: A survey with smartwatch users," in *VIS'20*, 2020, pp. 156–160.
- [3] S. Kashanj, X. Wang, and C. Perin, "Visualizations on smart watches while running: It actually helps!" in *VIS'24*, 2024, pp. 41–45.
- [4] J. Marshall, A. Dancu, and F. F. Mueller, "Interaction in motion: Designing truly mobile interaction," in *Proc. DIS'16*. ACM, 2016, p. 215–228. [Online]. Available: <https://doi.org/10.1145/2901790.2901844>
- [5] L. Pandey, K. Hasan, and A. S. Arif, "Acceptability of speech and silent speech input methods in private and public," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: <https://doi.org/10.1145/3411764.3445430>
- [6] J. Nie, Y. Fan, M. Zhao, R. Wan, Z. Xuan, M. Preindl, and X. Jiang, "Multi-modal dataset across exertion levels: Capturing post-exercise speech, breathing, and phonocardiogram," in *Proceedings of the 23rd ACM Conference on Embedded Networked Sensor Systems*, ser. SenSys '25. New York, NY, USA: Association for Computing Machinery, 2025, p. 297–304. [Online]. Available: <https://doi.org/10.1145/3715014.3722065>
- [7] M. Kytö, B. Ens, T. Piumsomboon, G. A. Lee, and M. Billinghurst, "Pinpointing: Precise head- and eye-based target selection for augmented reality," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–14. [Online]. Available: <https://doi.org/10.1145/3173574.3173655>
- [8] M. Hosseini, T. Ihmels, Z. Chen, M. Koelle, H. Müller, and S. Boll, "Towards a consensus gesture set: A survey of mid-air gestures in hci for maximized agreement across domains," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: <https://doi.org/10.1145/3544548.3581420>
- [9] H. Istance and et al., "Snap clutch, a moded approach to solving the midas touch problem," 01 2008, pp. 221–228.
- [10] R. J. Teather and W. Stuerzlinger, "Pointing at 3D targets in a stereo head-tracked virtual environment," in *2011 IEEE Symposium on 3D User Interfaces (3DUI)*, Mar. 2011, pp. 87–94. [Online]. Available: <https://ieeexplore.ieee.org/document/5759222?arnumber=5759222>
- [11] T. Luong, Y. F. Cheng, M. Möbus, A. Fender, and C. Holz, "Controllers or bare hands? a controlled evaluation of input techniques on interaction performance and exertion in virtual reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 11, pp. 4633–4643, 2023.
- [12] M. Amini, W. Stuerzlinger, R. J. Teather, and A. U. Batmaz, "A Systematic Review of Fitts' Law in 3D Extended Reality," in *CHI '25*, 2025.
- [13] I. S. MacKenzie, "Fitts' Law as a Research and Design Tool in Human-Computer Interaction," *Human-Computer Interaction*, vol. 7, no. 1, pp. 91–139, Mar. 1992, publisher: Taylor & Francis. eprint: https://doi.org/10.1207/s15327051hci0701_3. [Online]. Available: https://doi.org/10.1207/s15327051hci0701_3

- [14] Q. Zhou, D. Yu, M. N. Reinoso, J. Newn, J. Goncalves, and E. Velloso, "Eyes-free target acquisition during walking in immersive mixed reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 12, pp. 3423–3433, 2020.
- [15] Y. Lu, B. Gao, H. Tu, H. Wu, W. Xin, H. Cui, W. Luo, and H. B.-L. Duh, "Effects of physical walking on eyes-engaged target selection with ray-casting pointing in virtual reality," *Virtual Real.*, vol. 27, no. 2, p. 603–625, Aug. 2022. [Online]. Available: <https://doi.org/10.1007/s10055-022-00677-9>
- [16] Y. Li, J. Liu, J. Huang, Y. Zhang, X. Peng, Y. Bian, and F. Tian, "Evaluating the effects of user motion and viewing mode on target selection in augmented reality," *International Journal of Human-Computer Studies*, vol. 191, p. 103327, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1071581924001113>
- [17] T. Li, E. Velloso, A. Withana, and Z. Sarsenbayeva, "Estimating the Effects of Encumbrance and Walking on Mixed Reality Interaction," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, ser. CHI '25. New York, NY, USA: Association for Computing Machinery, Apr. 2025, pp. 1–24. [Online]. Available: <https://dl.acm.org/doi/10.1145/3706598.3713492>
- [18] Y. Shin, A. Esteves, and I. Oakley, "Using augmented reality on the go: Understanding the effects of mobility on user performance and subjective workload across eye, head, and hand ray pointing," *International Journal of Human-Computer Studies*, vol. 204, p. 103597, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1071581925001545>
- [19] I. O. f. S. , "ISO/TS 9241-411:2014 Ergonomics of human-system interaction. Evaluation methods for the design of physical input devices," 2014. [Online]. Available: <https://bsol-bsigroup-com.ap1.proxy.openathens.net/Bibliographic/BibliographicInfoData/00000000030298576>
- [20] T. Dwyer, K. Marriott, T. Isenberg, K. Klein, N. Riche, F. Schreiber, W. Stuerzlinger, and B. H. Thomas, *Immersive Analytics: An Introduction*. Cham: Springer International Publishing, 2018, pp. 1–23. [Online]. Available: https://doi.org/10.1007/978-3-030-01388-2_1
- [21] B. Spittle, M. Frutos-Pascual, C. Creed, and I. Williams, "A review of interaction techniques for immersive environments," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 9, pp. 3900–3921, 2023.
- [22] A. U. Batmaz, A. K. Mutasim, M. Malekmakan, E. Sadr, and W. Stuerzlinger, "Touch the Wall: Comparison of Virtual and Augmented Reality with Conventional 2D Screen Eye-Hand Coordination Training Systems," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Mar. 2020, pp. 184–193, iSSN: 2642-5254. [Online]. Available: <https://ieeexplore.ieee.org/document/9089504>
- [23] J. Liu, B. Ens, A. Prouzeau, J. Smiley, I. K. Nixon, S. Goodwin, and T. Dwyer, "Datadancing: An exploration of the design space for visualisation view management for 3d surfaces and spaces," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: <https://doi.org/10.1145/3544548.3580827>
- [24] F. Lu and et al., "Glanceable ar: Evaluating information access methods for head-worn augmented reality," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2020, pp. 930–939.
- [25] M. Khamis, C. Oechsner, F. Alt, and A. Bulling, "Vrpursuits: interaction in virtual reality using smooth pursuit eye movements," in *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, ser. AVI '18. New York, NY, USA: Association for Computing Machinery, 2018. [Online]. Available: <https://doi.org/10.1145/3206505.3206522>
- [26] T. Kosch, A. Matvienko, F. Müller, J. Bersch, C. Katins, D. Schön, and M. Mühlhäuser, "Notibike: Assessing target selection techniques for cyclist notifications in augmented reality," *Proc. ACM Hum.-Comput. Interact.*, vol. 6, no. MHCI, Sep. 2022. [Online]. Available: <https://doi.org/10.1145/3546732>
- [27] A. Michael and et al., "Race Yourself: A Longitudinal Exploration of Self-Competition between Past, Present, and Future Performances in a VR Exergame," in *Conference on Human Factors in Computing Systems - Proc.* ACM, Apr. 2020.
- [28] F. Grioui and et al., "Study of heart rate visualizations on a virtual smartwatch," in *Proc. of the 27th ACM Symposium on Virtual Reality Software and Technology*. ACM, 2021. [Online]. Available: <https://doi.org/10.1145/3489849.3489913>
- [29] E. Wu, T. Nozawa, F. Perteneder, and H. Koike, "Vr alpine ski training augmentation using visual cues of leading skier," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2020, pp. 3836–3845.
- [30] T. Matsumoto, E. Wu, and H. Koike, "Skiing, fast and slow: Evaluation of time distortion for vr ski training," in *Proceedings of the Augmented Humans International Conference 2022*, ser. AHs '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 142–151. [Online]. Available: <https://doi.org/10.1145/3519391.3519402>
- [31] T. Lin and et al., "Towards an understanding of situated ar visualization for basketball free-throw training," in *arXiv.org*. Cornell University Library, arXiv.org, 2021.
- [32] Y. Wu and et al., "AR-Enhanced Workouts: Exploring Visual Cues for At-Home Workout Videos in AR Environment," in *Proc. of the 36th Annual ACM Symposium on User Interface Software and Technology*. ACM, Oct. 2023, pp. 1–15. [Online]. Available: <https://dl.acm.org/doi/10.1145/3586183.3606796>
- [33] C. Ioannou, P. Archard, E. O'Neill, and C. Lutteroth, "Virtual performance augmentation in an immersive jump & run exergame," in *Proc. CHI'19*. ACM, 2019, p. 1–15. [Online]. Available: <https://doi.org/10.1145/3290605.3300388>
- [34] T. Hamada, A. Hautasaari, M. Kitazaki, and N. Koshizuka, "Solitary jogging with a virtual runner using smartglasses," in *IEEE VR'22*, 2022, pp. 644–654.
- [35] A. Li, C. Perin, J. Knibbe, G. Demartini, S. Viller, and M. Cordeil, "Embedded and situated visualisation in mixed reality to support interval running," *Computer graphics forum*, 2025.
- [36] J. Lee, S. C. Ahn, and J.-I. Hwang, "A walking-in-place method for virtual reality using position and orientation tracking," *Sensors*, vol. 18, no. 9, 2018. [Online]. Available: <https://www.mdpi.com/1424-8220/18/9/2832>
- [37] M. Hedlund, A. Lundström, C. Bogdan, and A. Matvienko, "Jogging-in-place: Exploring body-steering methods for jogging in virtual environments," in *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia*, ser. MUM '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 377–385. [Online]. Available: <https://doi.org/10.1145/3626705.3627778>
- [38] F. Müller, M. Schmitz, D. Schmitt, S. Günther, M. Funk, and M. Mühlhäuser, "Walk the line: Leveraging lateral shifts of the walking path as an input modality for head-mounted displays," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–15. [Online]. Available: <https://doi.org/10.1145/3313831.3376852>
- [39] A. Sedighi, E. Rashedi, and M. A. Nussbaum, "A head-worn display ("smart glasses") has adverse impacts on the dynamics of lateral position control during gait," *Gait & Posture*, vol. 81, pp. 126–130, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0966636220302551>
- [40] C. Zhou, Z. Yan, A. Ram, Y. Gu, Y. Xiang, C. Liu, Y. Huang, W. T. Ooi, and S. Zhao, "Glassmail: Towards personalised wearable assistant for on-the-go email creation on smart glasses," in *Proceedings of the 2024 ACM Designing Interactive Systems Conference*, ser. DIS '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 372–390. [Online]. Available: <https://doi.org/10.1145/3643834.3660683>
- [41] D. Gallo and et al., "RunAhead: Exploring Head Scanning based Navigation for Runners," in *Proc. of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Apr. 2020, pp. 1–13. [Online]. Available: <https://dl.acm.org/doi/10.1145/3313831.3376828>
- [42] C. J. Lin and B. H. Woldegiorgis, "Egocentric distance perception and performance of direct pointing in stereoscopic displays," *Applied Ergonomics*, vol. 64, pp. 66–74, Oct. 2017. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0003687017301187>
- [43] M. D. Barrera Machuca and W. Stuerzlinger, "The Effect of Stereo Display Deficiencies on Virtual Hand Pointing," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow Scotland UK: ACM, May 2019, pp. 1–14. [Online]. Available: <https://dl.acm.org/doi/10.1145/3290605.3300437>
- [44] R. Shi, Y. Wei, Y. Li, L. Yu, and H.-N. Liang, "Expanding Targets in Virtual Reality Environments: A Fitts' Law Study," in *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, Oct. 2023, pp. 615–618, iSSN: 2771-1110. [Online]. Available: <https://ieeexplore.ieee.org/document/10322135?arnumber=10322135>
- [45] K. McAnally, K. Wallwork, and G. Wallis, "The efficiency of visually guided movement in real and virtual space," *Virtual Reality*, vol. 27, no. 2, pp. 1187–1197, Jun. 2023. [Online]. Available: <https://doi.org/10.1007/s10055-022-00724-5>
- [46] J.-H. Sünderkamp, H. Stefanidi, A. Meschtscherjakov, and M. Tatzgern, "From lab to sidewalk: Exploring ar interaction techniques outdoors and in motion using fitts' law," in *Proceedings of the 2025 31st ACM*

- Symposium on Virtual Reality Software and Technology*, ser. VRST '25. New York, NY, USA: Association for Computing Machinery, 2025. [Online]. Available: <https://doi.org/10.1145/3756884.3766006>
- [47] M. J. McGuffin and R. Balakrishnan, "Fitts' law and expanding targets: Experimental studies and designs for user interfaces," *ACM Trans. Comput.-Hum. Interact.*, vol. 12, no. 4, p. 388–422, Dec. 2005. [Online]. Available: <https://doi.org/10.1145/1121112.1121115>
- [48] P. Manakhov, L. Sidenmark, K. Pfeuffer, and H. Gellersen, "Filtering on the Go: Effect of Filters on Gaze Pointing Accuracy During Physical Locomotion in Extended Reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 30, no. 11, pp. 7234–7244, Nov. 2024. [Online]. Available: <https://ieeexplore.ieee.org/document/10672561>
- [49] D. Bowman, C. Wingrave, J. Campbell, and V. Ly, "Using pinch gloves for both natural and abstract interaction techniques in virtual environments," *Proceedings of HCI International*, 01 2001.
- [50] S. W. Ducharme, D. S. Turner, J. D. Pleuss, C. C. Moore, J. M. Schuna, C. Tudor-Locke, and E. J. Aguiar, "Using cadence to predict the walk-to-run transition in children and adolescents: A logistic regression approach," *Journal of sports sciences*, vol. 39, no. 9, pp. 1039–1045, 2021.
- [51] C. J. Chase, E. J. Aguiar, C. C. Moore, S. R. Chipkin, J. Staudenmayer, C. Tudor-Locke, and S. W. Ducharme, "Cadence (steps/min) as an indicator of the walk-to-run transition," *Human Movement Science*, vol. 90, p. 103117, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167945723000635>
- [52] R. S. Woodworth, "Accuracy of voluntary movement," *The Psychological Review: Monograph Supplements*, vol. 3, no. 3, p. i, 1899.
- [53] R. Pastel, "Investigating the difficulty of one degree of freedom positioning: Associating movement phases with regions," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 53, no. 15, pp. 965–969, 2009. [Online]. Available: <https://doi.org/10.1177/154193120905301506>
- [54] C.-L. Deng, P. Geng, Y.-F. Hu, and S.-G. Kuai, "Beyond fitts's law: A three-phase model predicts movement time to position an object in an immersive 3d virtual environment," *Human Factors*, vol. 61, no. 6, pp. 879–894, 2019. [Online]. Available: <https://doi.org/10.1177/0018720819831517>
- [55] D. E. Meyer, R. A. Abrams, S. Kornblum, C. E. Wright, and J. E. Keith Smith, "Optimality in human motor performance: Ideal control of rapid aimed movements," *Psychological review*, vol. 95, no. 3, pp. 340–370, 1988.
- [56] L. Liu, R. van Liere, C. Nieuwenhuizen, and J.-B. Martens, "Comparing aimed movements in the real world and in virtual reality," in *2009 IEEE Virtual Reality Conference*, 2009, pp. 219–222.
- [57] K. Nieuwenhuizen, L. Liu, R. van Liere, and J.-B. Martens, "Insights from dividing 3d goal-directed movements into meaningful phases," *IEEE Computer Graphics and Applications*, vol. 29, no. 6, pp. 44–53, 2009.
- [58] J. C. Gower, "Generalized procrustes analysis," *Psychometrika*, vol. 40, no. 1, p. 33–51, 1975.
- [59] D. Wolf, J. Gugenheimer, M. Combosch, and E. Rukzio, "Understanding the heisenberg effect of spatial interaction: A selection induced error for spatially tracked input devices," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–10. [Online]. Available: <https://doi.org/10.1145/3313831.3376876>
- [60] S. G. Hart and et al., "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," in *Human Mental Workload*. North-Holland, 1988, vol. 52, pp. 139–183. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0166411508623869>
- [61] G. A. BORG, "Psychophysical bases of perceived exertion," *Medicine and science in sports and exercise*, vol. 14, no. 5, pp. 377–381, 1982.
- [62] P. Royston, "Remark as r94: A remark on algorithm as 181: The w-test for normality," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 44, no. 4, pp. 547–551, 1995. [Online]. Available: <http://www.jstor.org/stable/2986146>
- [63] R. A. Fisher and F. Yates, *Statistical tables for biological, agricultural and medical research*. Edinburgh and London: Oliver & Boyd Ltd., 1963.
- [64] D. A. Winter, *Signal Processing*, 4th ed. Hoboken, NJ: John Wiley & Sons, 2009, ch. 2, pp. 14–44. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470549148.ch2>
- [65] L. Liu and R. van Liere, "Designing 3d selection techniques using ballistic and corrective movements," 01 2009, pp. 1–8.
- [66] R. J. Teather, A. Pavlovych, W. Stuerzlinger, and I. S. MacKenzie, "Effects of tracking technology, latency, and spatial jitter on object movement," in *2009 IEEE Symposium on 3D User Interfaces*, 2009, pp. 43–50.
- [67] Y. Li, B. Tag, S. Dai, R. Crowther, T. Dwyer, P. Irani, and B. Ens, "Nicer: A new and improved consumed endurance and recovery metric to quantify muscle fatigue of mid-air interactions," *ACM Trans. Graph.*, vol. 43, no. 4, Jul. 2024. [Online]. Available: <https://doi.org/10.1145/3658230>
- [68] H. Kim, Y. Hong, J. Yu, S. Xiong, and W. Kim, "Toward a more standardized multi-directional tapping task in vr: the effect of target depth," in *2025 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2025, pp. 359–369.
- [69] K. L. Havens, T. Mukherjee, and J. M. Finley, "Analysis of biases in dynamic margins of stability introduced by the use of simplified center of mass estimates during walking and turning," *Gait & Posture*, vol. 59, pp. 162–167, 2018.