Effective Throughput Analysis of Different Task Execution Strategies for Mid-Air Fitts' Tasks in Virtual Reality

Anil Ufuk Batmaz and Wolfgang Stuerzlinger

Abstract—Fitts' law and throughput based on effective measures are two mathematical models frequently used to analyze human motor performance in a standardized pointing task, e.g., to compare the performance of input and output devices. Even though pointing has been deeply studied in 2D, it is not well understood how different task execution strategies affect throughput in pointing in 3D virtual environments. In this work, we examine the effective throughput measure, claimed to be invariant to task execution strategies, in Virtual Reality (VR) systems with three such strategies, "as fast, as precise, and as fast and as precise as possible" for ray casting and virtual hand interaction, by re-analyzing data from a 3D pointing IS0 9241-411 study. Results show that effective throughput is not invariant for different task execution strategies in VR, which also matches a more recent 2D result. Normalized speed vs. accuracy curves also did not fit the data. We thus suggest that practitioners, developers, and researchers who use MacKenzie's effective throughput formulation should consider our findings when analyzing 3D user pointing performance in VR systems.

Index Terms—Fitts' task, Virtual Reality, Effective Throughput, Speed-Accuracy Trade-off

1 Introduction

The classic Fitts' law task of hitting alternating targets [28] has been used to analyze and assess user performance for pointing tasks in 2D and also 3D, e.g., [73]. With this methodology, various input devices, including mice [66], pens [74], styli [46] or laser pointers [55], have been investigated in combination with different output devices, such as monitors [71], large screens [42], mobile devices [49], physical apparatuses [4,28], or even CAVE systems [72]. Building on decades of research, MacKenzie's formulation of Fitts' law is the most frequently used variation in Human-Computer Interaction research, see, e.g., [47]:

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

In Equation 1, A is the target distance and W the target size. The *log* term represents the task difficulty or the *index of difficulty*, *ID*. The coefficients a and b are empirically derived via linear regression.

Apart from modeling the movement time, Fitts' law research also frequently uses a second mathematical model to analyze user performance, namely the index of performance. One of the most well-known versions and the one predominantly used in current Human-Computer Interaction studies, is MacKenzie's formulation for effective throughput [48], originally motivated via Shannon's channel capacity theorem.

$$Throughput = \left(\frac{ID_e}{MovementTime}\right) \qquad (2)$$

In Equation 2, movement time is the task execution time and ID_e the effective index of difficulty, with accounts for the effect of the combination of user's *accuracy and precision* in ISO pointing tasks [38]:

$$ID_e = \log_2\left(\frac{A_e}{W_e}+1\right) \qquad (3)$$

In Equation 3, the effective target distance A_e is the average real distance traversed with the cursor to execute the task and W_e is the effective target width, calculated as $W_e = 4.133 \times SD_x$, where SD_x is the standard deviation of the distance between the target center and the selection coordinates projected onto the task axis. I.e., this is a uni-variate formulation that does not account for deviations in the direction orthogonal to the main task direction. SD_x represents the **accuracy** of the task execution [38,50,51], i.e., how close the selection points are to the target center. The SD_x calculation is illustrated in Fig. 1, based on [49]. In this figure, d_x is the distance on the task selection axis and SD_x is the standard deviation in the d_x values. d_x is calculated as $d_x = (k^2 – j^2 – i^2)/(2 \times i)$. To maintain comparability with 2D work on the same topic, we used the uni-variate distribution of the selection points [81] to characterize accuracy, which also allowed us to compare our results directly with Mackenzie and Isokoski [50].

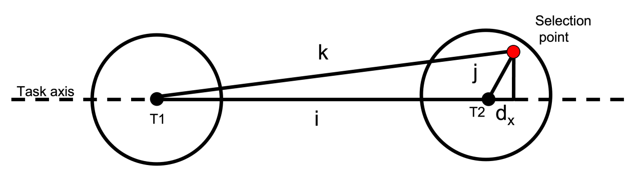


Fig. 1. Geometry to calculate the deviation d_x along the task axis [49].

The ISO standard presents Fitts' Law and the above-mentioned definition of effective throughput as an "evaluation method for the design of physical input devices for interactive systems" [38]. According to ISO 9241-411, ID_e represents the task **precision** or "measure of the user precision achieved in accomplishing a task" [38]. Precision here refers to how close the selection points are to each other.

Equation 1 encodes the relationship between task execution time, effective target size, and movement distance. Equation 2 then combines *time, precision,* and *accuracy* into *a single measure*, which enables comparison of different input and output devices, as well as human motor performance under different conditions.

Some studies of MacKenzie's throughput formulation questioned the validity of the proposed model in 2D. Guiard et al.'s work [33] demonstrated that throughput is systematically affected by task execution strategy. In follow-up work, [56] Olafsdottir's et al.'s also asked participants to execute the task with different execution strategies, e.g., "as fast as possible" or "as precise as possible". Their results identified that effective throughput depends on the task execution strategy. Both these results for the speed-accuracy trade-off do not match the findings of Mackenzie and Isokoski's earlier work [50]. Yet, all these works investigated only *2D pointing tasks*.

More than a decade ago, research in virtual reality (VR) interaction adopted the above-mentioned methodology to analyze mid-air user interaction performance in 3D virtual environments(VEs) [73]. Yet,

- *Anil Ufuk Batmaz is with the Computer Science Department at Concordia University. E-mail: ufuk.batmaz@concordia.ca.*
- *Wolfgang Stuerzlinger is with the School of Interactive Arts & Technology at Simon Fraser University. E-mail: w.s@sfu.ca.*

due to the additional degrees of freedom that 3D VEs offer, a selection task is more *challenging* with 3D mid-air controllers than with a 2D device. One issue is that the original Fitts' task involved tapping on table surfaces [28], which means that the lack of haptic feedback in mid-air might affect the outcome [72]. Additionally, while users interact with a mid-air 3D input device, such as a VR controller, the system is prone to jitter from different sources, including hand tremor and tracking system noise. Further, such jitter is amplified with distance in ray pointing [12], where, e.g., a small, $<5°$ movement of a VR controller moves the cursor 8 cm at 1 m, but 24 cm at 3 m. Thus, the location the user points at is (naturally) different for every selection, which, typically increases the SD_x used in the throughput analysis relative to 2D work. Additional factors, such as conflicting depth cues [4, 6], can also play a substantial role.

Given all these issues it is not guaranteed that Fitts' law is directly applicable for modeling 3D movements. Several prior studies thus proposed new versions of Fitts' law and verified that these mathematical models describe user movement times more accurately [4, 20, 21, 54]. However, these efforts focused solely on developing a user movement time model, ignoring the effect of user accuracy and precision. As such, these new models are beyond the scope of our current investigation, also because none of them defined an appropriate throughput formulation. Thus, we do not propose or aim to analyze a new model for 3D Fitts' law for mid-air interaction in this work.

As previous work demonstrated that effective throughput increases the comparability of results across studies [68], Fitts' law and effective throughput were used in 3D user interface studies for several purposes: to compare the performance of input devices, to analyze the speed-accuracy trade-off through the ISO9241-411 formula, and as a motor performance assessment criterion, e.g., [9,10,13,15,16,18,22,25,40,63, 73–75]. All these studies assume that effective throughput is invariant. Yet, to our knowledge, the claim of invariance of effective throughput has not been validated for mid-air interaction before, i.e., it is unknown if the 2D formulation is still valid and useful for systematic comparisons in 3D. Thus, researchers still rely on 2D effective throughput measures for 3D pointing experiments. Previous work already highlighted the corresponding need to investigate "the variance of effective throughput for task execution strategies in different systems" [33, 56]. To address this gap, our current work is the first that investigates *how throughput varies in 3D selection with different task execution strategies*.

To investigate the variation of user performance in VR with different task execution strategies, we analyzed data from a previous study [14] where participants had been asked to execute a 3D version of the ISO 9241-411 pointing task [38] "as fast as possible", "as precise as possible", or "as fast and as precise as possible".

That previous study [14] focused solely on the effects of auditory error feedback in 3D Fitts' law studies for VR training systems. While the findings, contributions, hypotheses, and research questions of their study 2 [14] investigate how user performance varies with different auditory error feedback under various task execution strategies, that work *did not analyze throughput in depth* . The outcomes of that work suggest that, to increase user performance, it is best to avoid using higher frequencies in auditory feedback. Still, this previous work did not investigate the variance of effective throughput with task execution strategy in 3D ISO 9241-411 tasks.

Through reanalyzing the data from this previous study, we examine the following hypothesis: **Effective throughput varies for different task execution strategies in mid-air pointing.** In other words, we do not expect to observe a throughput invariance with different task execution strategies, as claimed by Mackenzie and Isokoski [50], but to observe throughput variance, in line with other studies [33, 56].

We believe that the results presented here will affect how throughput is used in studies that use ISO 9241-411 to compare input devices and in VR training systems or simulators.

2 Previous Work

2.1 Mid-Air Interaction Studies with Effective Throughput

Interaction with a virtual hand and ray casting are the two most frequently used mid-air 3D interaction techniques in VR [2, 44]. While the virtual hand technique is mostly used for peri-personal space interaction, i.e., objects within arm's reach, up to approximately 70 cm from the user, ray casting enables users to also interact with virtual objects beyond arm's length [4, 10, 25]. Both techniques are widely used as baselines in VR and Human-Computer Interaction studies and frequently compared in VR systems [26].

Mid-air interaction in immersive systems can differ from interaction in the 2D plane, e.g., on desktop monitors, as targets can be arranged in 3D. Still, much work on 3D pointing has used planar arrangements of targets. If the plane on which the targets are faces the user, i.e., if it is orthogonal to the view direction, the task is (generally) equivalent to a 2D Fitts' law task and Fitts' law holds well, which also increases the comparability with 2D work. If the (visual) depth, i.e., the distance from the user, changes substantially between subsequent targets, research for 3D pointing with the virtual hand technique has shown that the 2D formulation of Fitts' law does not hold, potentially due to the effect of conflicting depth cues [4]. Yet, this is still an active area of ongoing research [20, 21, 54, 73].

2.2 Fitts' Law and Throughput as Assessment Criterion

Previous work showed that Fitts' law and throughput based on effective measures can be used to assess human performance with different input devices in VR. For instance, Teather and Stuerzlinger [74] examined participants' pointing performance in a fish-tank VR system with this measure. Recent studies also used throughput to analyze user performance in immersive VR training systems and simulators [9, 18, 77].

Batmaz et al. identified that individuals can benefit from prioritizing different criteria when they learn a new task and that novices should prioritize precision over speed to improve their learning curve [7]. This was achieved through active instruction, i.e., telling participant to be faster or more precise [7].

3 User Study

For the main analysis, we use a subset of data from previously published work [14]. Such re-use of data [14] is frequent in various research fields, e.g., [27, 35, 78, 79]. Although we use data from a prior study, we aim to keep the current paper self-contained to facilitate reader understanding. Thus, we also explain all relevant details of the user study that was used to collect the data, specifically the participant characteristics and the procedure. Since we only re-analyze a subset of the data here, there were no changes to the apparatus or procedure. Thus, all the information about the apparatus and procedure for the user study is the same as in previous work [14]. The ethics approval for the original study also permitted the reuse of the data collected from the participants.

The subset of the data used here was collected in a mid-air pointing experiment where the participants were asked to focus on different task execution strategies with different *auditory error feedback* conditions [14]. For the current paper, and after we had verified that the subset was not affected by ordering effects (see the supplemental material) we only used the data from the *constant auditory feedback* condition of that experiment, i.e., the condition that matches previous work on throughput in terms of feedback.

3.1 Participants

18 participants took part in the study (17 right-handed and one left-handed, 10 male and 8 female) with an average age of 29.31 ± 4.29. The experiment was conducted remotely and data collected only from participants that had a computer able to run Steam VR under Windows 10. Following the suggestions of Steed et al. for remote studies [70], headsets were running with at least 90 Hz. Seven participants used an HTC Vive Pro, six a HTC Vive, three an Oculus Quest, and two participants used an Oculus Rift. Eleven participants reported that they play computer games 0-5 hours weekly, four 5-10 hours, and three participants 10-20 hours. Eleven participants reported that they use 3D CAD systems between 0-5 hours weekly and seven 5-10 hours. Each individual was asked to adjust the inter-pupillary distance of their headset before the experiment.

3.2 Procedure

Participants started the experiment by filling a demographic questionnaire, followed by an explanation of the procedure. In the virtual environment, participants were sitting in the middle of an empty room with pictorial depth cues, as shown in Fig. 2.

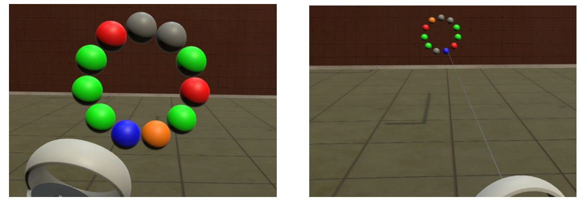
The experiment used an ISO 9241-411:2012 [38] task with 11 gray targets placed in a circle, with the circle center positioned at the participants' eye level. The first target was randomly chosen by the system and participants executed the task either in a clockwise or counter-clockwise direction, also randomly chosen by the software. During task execution, and while the cursor was inside the target, we changed the color of the sphere temporarily to blue for visual feedback. If the cursor was inside the orange-colored, i.e., desired, target when selection occurred, the target's color was changed to green to provide positive feedback to the participant. On the other hand, when the cursor was outside the target during selection, the target's color was changed to red and a sound played to provide error feedback. During the experiment, participants used their dominant hand to control the cursor with the VR controller. To avoid the negative consequences of the Heisenberg effect [17], i.e., where the user applies a mechanical force on the button to select a target and that force moved the physical controller and thus the cursor, we instead asked participants to press the space button on a keyboard with their non-dominant hand to indicate selection.

Targets were placed either 0.4 or 1.5 m away from the subjects for the virtual hand and ray casting conditions, respectively, and participants could easily see all targets in their field of view. These distances match the task IDs across the two input conditions and are also similar to previous work on mid-air selection [10]. Participants selected each target with the cursor associated with the VR controller with two different **selection techniques**, virtual hand (Fig. 2(a)) and ray casting (Fig. 2(b)). To eliminate diplopia, i.e., potential double-vision in the virtual hand condition, the 1 cm "cursor" sphere was placed 3 cm above the VR controller. In the ray casting condition, the cursor at the end of the ray was always in the same plane where targets were positioned, i.e., 1.5 m away from the user. Participants selected targets at 0.4 m with the virtual hand technique and at 1.5 m with the ray casting technique.

Subjects performed the experiment with three different **task execution strategies**. In the first condition, participants were asked to perform the experiment "as fast as possible," i.e., to perform the task while focusing only on their task execution time. In the second one, subjects were asked to perform the experiment "as precisely as possible", i.e., to focus only on precisely selecting the targets while ignoring speed. As the third condition, participants were asked to perform the experiment "as fast and as precise as possible", i.e., to focus both on their speed and precision simultaneously. All previously mentioned motor performance training studies used one or more of these three task execution strategies, which makes our choice of task execution strategies directly relevant to motor performance training [1, 23, 24, 29, 36, 37, 43, 58, 59, 61, 65, 69]. Finally, instructions for the current **task execution strategy** were verbally stated by the experimenter at the beginning of the each round of pointing trials for each condition.

To further assist participants in keeping track of the current strategy, the current task execution strategy was shown as floating text behind the targets during the experiment. The experimenter also monitored each participants' performance over a video link and tried to ensure that participants followed the current task execution strategy through spoken feedback. For instance, if a subject was repeatedly making errors in the "as precise as possible" condition, the experimenter encouraged the participant to slow down and to focus on precision.

To vary the index of difficulty (ID), three different **target sizes** (1.5, 2.5, and 3.5 cm) and two different **target distances** (12.5 and 25 cm) were used, resulting in 6 unique IDs between 2.19 and 4.14. These target sizes and distances match other 3D pointing work [4, 8] and follow Gori et al.'s study [32], who recommended linearly increasing sizes. To cancel potential learning effects, the **selection technique** and **task execution strategy** factors were counterbalanced with a Latin square design.



(a) (b)

Fig. 2. Illustration of task color scheme and selection techniques. The current target is shown in orange, blue is used for highlighting the target that the cursor currently collides with. A correct selection then counts as a "hit" and is shown in green, while "misses" are shown in red. Instances when the cursor is inside the target with the (a) virtual hand and (b) ray casting conditions.

3.3 (Remaining) Experimental Design

As mentioned above, the previous work used different auditory feedback conditions [14], but we investigate here only the data for the *constant auditory feedback* condition from that experiment. Consequently, we analyzed all dependent variables only for the remaining data (and thus the subset of the experimental design in terms of independent variables) to investigate the impact of effective throughput with an ISO 9241-411 task. For the full previous experimental design [14], please refer to the supplementary material.

For the data analyzed here, we effectively looked at a two-factor within-subjects design with three different **task execution strategies** (3_{TES} = as fast as possible, as precise as possible, and as fast and as precise as possible) and two **selection techniques** (2_{ST} = virtual hand and ray casting), comprising a $3_{TES} \times 2_{ST}$ design.

We analyzed the measured task execution time (seconds), error rate (%), and effective throughput (bits/s) of the participants. As previous work has shown that, compared to A_e, effective target size has a more significant effect on the throughput results [45], we also analyzed the SD_x data to investigate accuracy. Further, to ensure comparability with previous work [50, 56] and when analyzing the data for effective distance and target width, we used the distance between the selection point and the target as measured within the target plane. Other VR studies had used the same method to compute throughput, e.g., [25, 74].

In total, we analyzed the data of $3_{TES} \times 2_{ST} \times 6_{ID} \times$ 11 repetitions = 396 trials for each subject.

4 Results

The data for the (reduced) experimental design was analyzed using two-way repeated measures (RM) ANOVA in SPSS 24. We used Skewness (S) and Kurtosis (K) for normality analysis and considered data as normally distributed when the S and K values were within $\pm$ 1.5 [34, 52]. When the data was not normally distributed, we used ART [80]. We applied Huynh-Feldt correction when the ε was less than 0.75. *For brevity, we only report significant results.* We used the Bonferroni method for post-hoc analyses. We first analyzed the results of the two-way RM ANOVA, followed by the separate interaction techniques, and finally the one-way interactions.

4.1 Two-way interactions

We found significant interactions for time (F(2,34) = 17.879, $p < 0.001$, $\eta^2 = 0.513$), throughput (F(2,34) = 16.066, $p < 0.001$, $\eta^2 = 0.486$), and standard deviation of selection points (F(2,34) = 29.6, $p < 0.001$, $\eta^2 = 0.635$) between task execution strategy and selection techniques. These results are shown in Fig. 3(a), (b), and (c). We further analyzed the throughput results with the effective target distance (A_e) and did not find a significant interaction (S = -0.05, K = -1.96, F(2,34) = 1.002, p = 0.378, $\eta^2 = 0.056$), as shown in Fig. 3(d). According to these results, subjects were significantly faster and their throughput was significantly higher with the virtual hand selection technique with each task execution strategy. However, when they executed the task "as precise as possible", we did not observe a significant difference between ray casting and virtual hand in terms of accuracy.

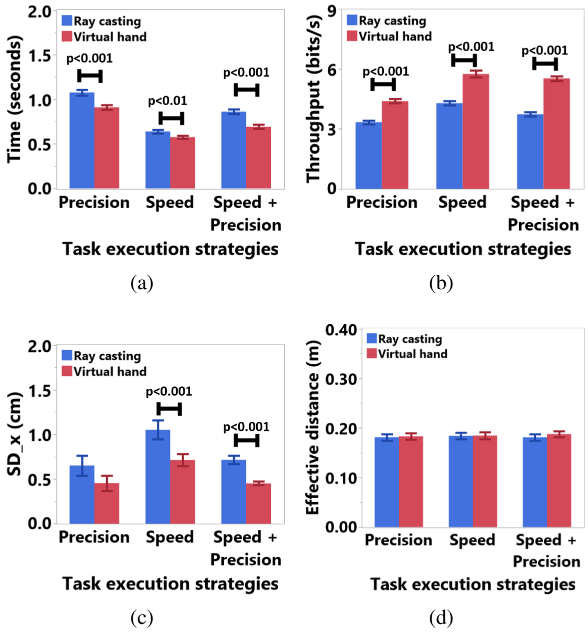


Fig. 3. Interaction of task execution strategy and selection technique for (a) time, (b) throughput, (c) accuracy and (d) effective target distance.

Detailed Analysis per Selection Technique In a detailed analysis for the virtual hand interaction technique, we looked only at the data for that condition and further analyzed the effect of task execution strategy. For virtal hand, throughput (S = 0.41, K = 0.11) was normally distributed. Time (S = -0.27, K = 0.27) and SD_x (S = 0.3, K = -0.6) were normal after log-transformation. We found significant interactions of the task execution strategy on time (F(2,34) = 159.76, $p < 0.001$, $\eta^2 = 0.904$), error rate (F(2,34) = 64.784, $p < 0.001$, $\eta^2 = 0.811$), throughput (F(2,34) = 73.958, $p < 0.001$, $\eta^2 = 0.813$), and SD_x (F(2,34) = 82.114, $p < 0.001$, $\eta^2 = 0.868$). These results are shown in Fig. 4.

Focusing only on the data for the ray casting technique, error rate (S = 0.98, K = 0.09) and throughput (S = 0.07, K = -0.19) were normally distributed. Time (S = 0.17, K = -0.46) and SD_x (S = -0.05, K = -0.35) were normal after log-transformation. Here we also found significant interactions of task execution strategy on time (F(2,34) = 76.00, $p < 0.001$, $\eta^2 = 0.817$), error rate (F(2,34) = 54.967, $p < 0.001$, $\eta^2 = 0.764$), throughput (F(2,34) = 45.58, $p < 0.001$, $\eta^2 = 0.728$), and SD_x (F(2,34) = 53.214, $p < 0.001$, $\eta^2 = 0.758$). These results are shown in Fig. 5.

According to detailed analysis results for each interaction technique, participants exhibited a higher throughput with the "as fast as possible" task execution strategy. The rest of the results are similar to the findings shown in Fig. 6.

4.2 One-way interactions

The results of the one-way ANOVAs are illustrated in Fig. 6 as means and standard error of means, with the statistics listed in a compact form in Table 1.

Time The dependent variable time was not normally distributed (S = 1.18, K = 1.83). According to the results in Table 1 and Fig. 6(a), participants were significantly faster with the "as fast as possible" task execution strategy. Moreover, individuals were significantly faster with the virtual hand technique, Fig. 6(e).

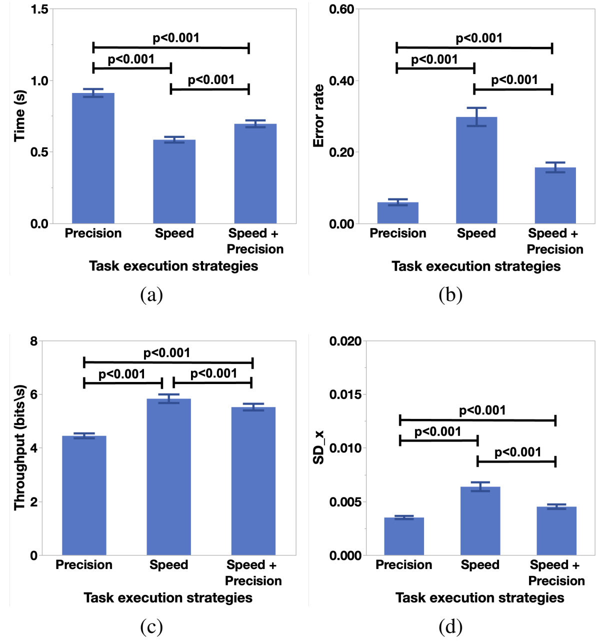


Fig. 4. Detailed task execution strategies results for virtual hand interaction technique: a) time, b) error rate c) throughput, and d) SD_x.

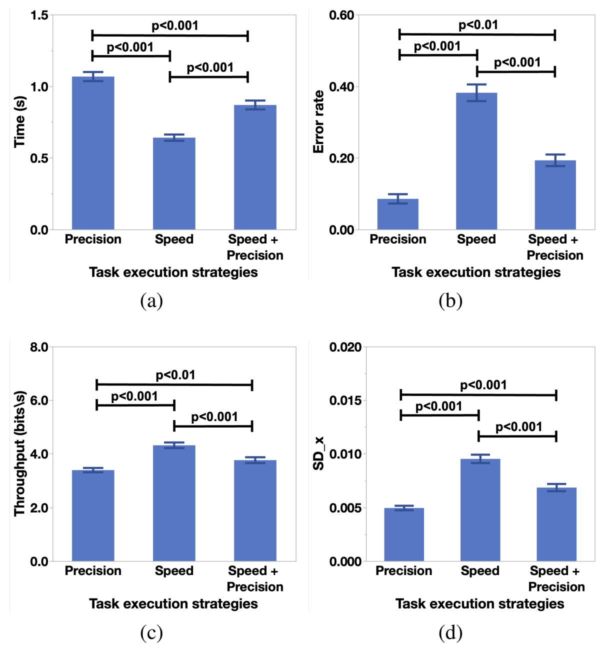


Fig. 5. Detailed task execution strategies results for ray casting interaction technique: a) time, b) error Rate c) throughput, and d) SD_x.

Table 1. One-way RM ANOVA results

	Task Execution Strategy	Selection Technique	Index of Difficulty
Time	F(2, 34) = 143.654, $p < 0.001$, $\eta^2 = 0.894$	F(1, 17) = 49.697, $p < 0.001$, $\eta^2 = 0.745$	F(5, 85) = 138.989, $p < 0.001$, $\eta^2 = 0.891$
Error rate	F(1.47, 24.95) = 76.876, $p < 0.001$, $\eta^2 = 0.819$	F(1, 17) = 14.618, $p < 0.001$, $\eta^2 = 0.462$	F(5, 85) =76.418, $p < 0.001$, $\eta^2 = 0.818$
Throughput	F(2, 34) = 72.271, $p < 0.001$, $\eta^2 = 0.810$	F(1, 17) =155.945, $p < 0.001$, $\eta^2 = 0.902$	F(3.711, 63.084) =22.816, $p < 0.001$, $\eta^2 = 0.573$
SD_x	F(2, 34) = 239.694, $p < 0.001$, η^2= 0.934	F(1, 17) = 239.694, $p < 0.001$, $\eta^2 = 0.934$	F(5, 85) = 44.306, $p < 0.001$, $\eta^2 = 0.723$

Error Rate The error rate variable was normally distributed (S = 1.19, K = 0.97). According to the results in Table 1, participants made significantly more errors with the "as fast as possible" task execution strategy and fewer errors with "as precise as possible," Fig. 6(b). Moreover, participants made significantly fewer errors with the virtual hand selection technique, Fig. 6(f).

Throughput The throughput variable was normally distributed (S = 0.33, K = 0.46). According to the results in Table 1 and Fig. 6(c), we observed a significant difference between task execution strategies, with significantly higher throughput appearing for the "as fast as possible" task execution strategy. Furthermore, subjects' throughput significantly increased with the virtual hand selection technique, Fig. 6(g).

Standard deviation SD_x The standard deviation dependent variable was not normally distributed (S = 3.31, K = 18.95). According to the results in Table 1, subjects' accuracy significantly decreased with the "as fast as possible" task execution strategy as seen in Fig. 6(d). Moreover, participants' accuracy significantly increased with the virtual hand condition, Fig. 6(h).

4.3 Detailed Speed-Accuracy Trade-off analysis

We analyzed the data in the same way as MacKenzie and Isokoski [50]. We first normalized the SD_x and movement time measurements and then plotted the data for all participants. As in MacKenzie and Isokoski's Figure 8 [50], and for clarity, we plot only the data points for the speed-emphasis and precision-emphasis conditions, see Fig. 7.

In Fig. 7, the blue curve shows constant throughput across different movement times and SD_xs. The figure also illustrates that, e.g., with an increase in movement time by 20%, the SD_x needs to drop by approximately 40% to yield the same throughput.

The non-linear regression results for Fig. 7 identified an R^2 value of 0.26 for the fit. Separate fits for ray casting and virtual hand conditions exhibit $R^2 = 0.48$ and $R^2 = 0.45$, respectively. Unlike the data shown in Figure 8 of the MacKenzie and Isokoski [50], it is thus not possible to observe throughput invariance in our data, neither visually nor quantitatively.

4.4 Fitts' Law analysis

The results for the index of difficulty (ID) in Table 1 show that time, error rate, throughput, and SD_x all significantly vary with the task difficulty. When we fit the data for time with Equation 1, the linear regression results showed that the "as precise as possible" task execution strategy can be modelled as MT = 0.08 + 0.82 * ID, $R^2 = 0.89$, "as fast as possible" as MT = 0.07 + 0.16 * ID, $R^2 = 0.96$, and the "as fast and as precise as possible" task execution strategy can be modelled as MT = 0.08 + 0.22 * ID, $R^2 = 0.91$. These results are shown in Fig. 8(a). Similarly, the results for ray casting and virtual hand selection techniques are shown in Fig. 8(b), with MT = 0.07 + 0.24 * ID, $R^2 = 0.91$ and MT = 0.11 + 0.19 * ID, $R^2 = 0.92$, respectively. As can be observed from Fig. 8, the task execution time increases with higher task difficulty.

5 Discussion

In this work, we analyzed 3D mid-air selection performance with a Fitts' ISO 9241 [38] task with three different task execution strategies: as fast, as precise, and as fast and as precise as possible.

Before analyzing the results, we empirically verified that participants were successfully following the task execution instructions. When we looked at data for individual participants, we did not observe strong outliers or other indications that behaviours around speed-accuracy trade-off vary fundamentally across people. Based on the significant differences between conditions, it seems that the participants readily understood the difference in the execution strategies. Our results confirm that subjects were indeed faster with "as fast as possible", while they were slowest with the "as precise as possible" task execution strategy. The error rate results similarly supports this observation. The time results in Fig. 4 and Fig. 5 also show that participants executed the task based on the given feedback.

The results for our study show a higher error rate for the "as fast as possible" execution strategy. We also observe that the SD_x results are

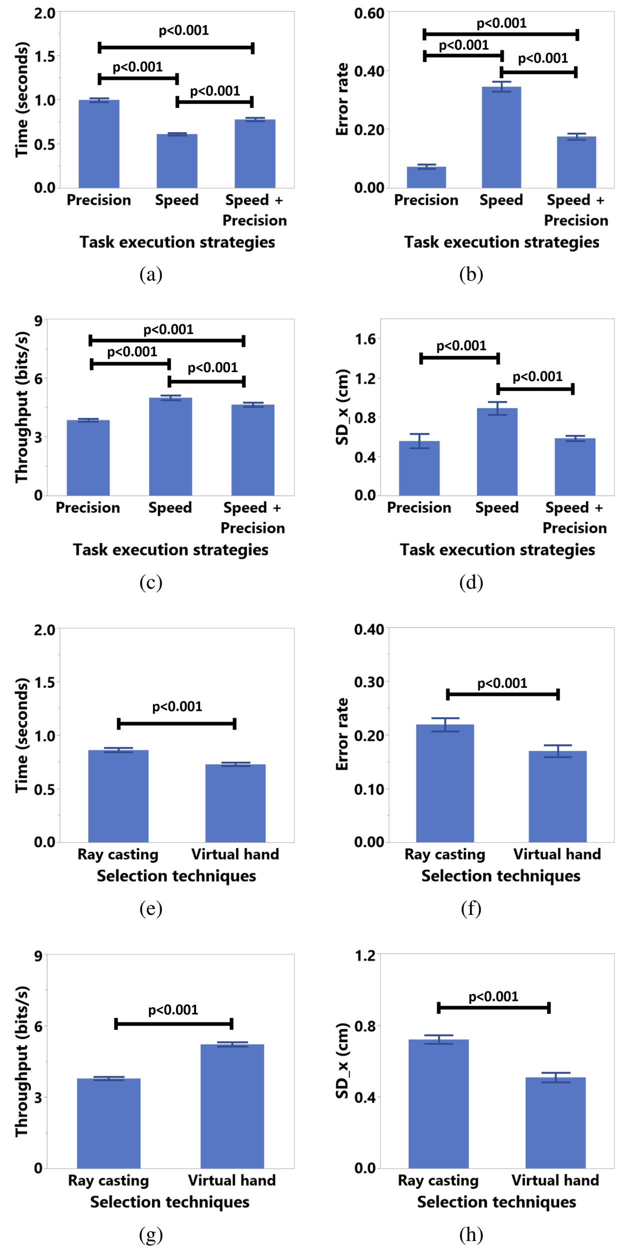


Fig. 6. Task execution strategy results for (a) time, (b) error rate, (c) throughput, and (d) accuracy. Selection technique results for (e) time, (f) error rate, (g) throughput, and (h) accuracy.

higher for both interaction techniques with this strategy. Even though one could expect confounding results for these dependent variables, Mackenzie and Isokoski already highlighted that *"it is well known that behaviour is more erratic when humans act with haste [41]"* [50].

From a statistical point of view, one could argue that the invariance of the throughput hypothesis can always be shown to be false; a big enough sample size could almost always lead to statistically significant differences [53]. However, the study did not feature a large number of participants nor a large number of pointing motions (18 participants, 7128 pointing trials). In comparison, Olasdottir et al. [56] used 16 participants and 6000 trials. Similarly, in MacKenzie's effective throughput study [50], there were 18 participants and 5400 trials. This means that the studies have comparable number of participants and trials. Furthermore, our results for effective throughput on task execution strategy exhibit a high effect size, which indicates a strong effect and thus also that the probability of arriving at the same result

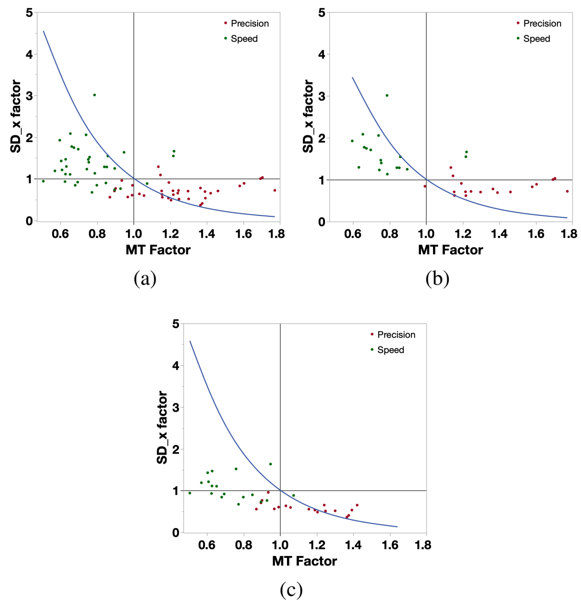


Fig. 7. Changes in speed (relative MT) and accuracy (relative SDx) with constant throughput for all participants, across a) all conditions, b) only ray casting and c) only virtual hand.

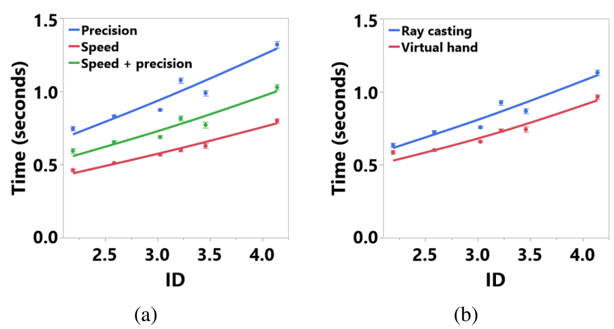


Fig. 8. Fitts' Law results for movement time for (a) task execution strategy and (b) interaction technique.

in a replication study is high. In addition to the statistical results, we also point to the relative speed-accuracy trade-off plots in Fig. 7, as these figures directly illustrate that throughput changes with different task execution strategies. The R^2 values for each interaction condition across participants are also low.

Our results match outcomes from previous work [10, 25] in that subjects are faster, more precise, and exhibit higher throughput with the virtual hand. Since the objects were closer to the user with the virtual hand, one might conclude that it was easier to select targets, but we note that the IDs for both conditions were the same. We believe that the reason behind this outcome is that ray casting is more prone to (rotational) jitter [12], which can negatively affect user performance in terms of time, error rate, and throughput. Furthermore, we also speculate that the conflicting depth depth cues in current VR HMDs decreased the interaction performance, as already identified in previous work [4, 6]. In additional, more detailed analysis for each interaction technique, we were unable to identify information that had "disappeared" in the averaging.

Switching to angular width and amplitude measures naturally changes the data shown in Fig. 8. Yet, after converting the widths and distances to the angular domain based on [42], we still see an ID range of 2.2-4.1, similar to previous work, and also a similar difference between interaction techniques. Thus, our findings on task execution strategies for mid-air pointing seem to hold regardless if we run the statistics via linear or angular measures. Still, we deliberately chose not to use the angular formulation to keep the comparability with 2D work on the same topic high.

5.1 Outcomes and Potential Impact of the Work

The results of our work point to three major outcomes. The first one is that our results contribute to the discussion around MacKenzie's [50] and Olafdottir's [56] results with a study involving mid-air pointing. Our work thus also extends the psycho-motor literature for Fitts' law. The second potential impact concerns the use of ISO 9241-411 in a VR study, to analyze user performance with physical input devices in interactive systems. The third outcome concerns using the ISO 9241-411 task within VR motor performance training systems. We discuss the details for each of the three main direct outcomes of our work for VR applications in the following subsections.

5.1.1 Comparison with Mackenzie's and Olafsdottir's results

MacKenzie and Isokoski's [50] and Olafsdottir et al.'s [56] previous work both used 2D setups with 2D input devices in their experiments. As mentioned before, their results cannot be assumed to automatically hold for mid-air VR systems. When we analyzed the results for throughput, we found that the "as fast as possible" task execution strategy had the highest throughput results for mid-air selection. Yet, unlike MacKenzie and Isokoski's findings [50], our results are in line with Olafsdottir et al.'s previous work [56]. Moreover, the results in Fig. 7 show that participants' performance does not follow a constant speed-accuracy trade-off curve. The variability in the speed-accuracy trade-off can be visually observed through the lack of a match between the data and the curve in Fig. 7 and through the low R^2 values. These findings, i.e., that **throughput performance is task execution dependent for mid-air pointing**, support our hypothesis.

Overall, our results show that MacKenzie's effective throughput model (Equation 3) depends on the task execution strategy for VR mid-air interaction methods, i.e., *it is not invariant*. When we analyzed virtual hand and ray casting technique results separately to investigate differences in the influence of the interaction technique on throughput invariance, we did not find significant results. Yet, if the participants executed the task precisely with ray casting, they were able to reach the same level of accuracy as with the virtual hand technique; however, this took more time and the individuals' performance in terms of throughput decreased. This is where our results clearly demonstrate the presence of a speed-accuracy trade-off in 3D pointing. On the other hand, and matching previous outcomes for 2D pointing [56], our results confirm that this issue does not pose an obstacle for the use of effective throughput to evaluate the design of novel input devices in 3D VR systems. However, the question of the (in)variance of effective throughput still requires more research, at least for 3D mid-air pointing.

5.1.2 Implications for ISO 9241-411 mid-air pointing studies

Our work also questions the application of ISO 9241-411 standards for effective throughput assessment in mid-air interaction. The results of our work support the findings of Olasdottir et al. [56] and Gori et al. [32] in terms of the (effective) throughput being task execution strategy dependent. However, our findings do not reveal the possible cause for the difference with Mackenzie and Isokoski's results [50]. One potential explanation for the dependency of the throughput measurement on task execution strategy in 3D mid-air pointing could be the technical limitations of the VR systems, such as jitter [11, 12], which means that future research is needed to explore this issue.

Participants still had to position the VR controller in (visual) depth to select targets in the virtual hand condition in the ISO 9241-411 task. Similarly, with the ray casting method, all controller depth movements had also a potential impact on the position of the cursor. Even though the task here used a 2D planar arrangement of targets facing the user, i.e., a task where pointing is typically well-modeled by the 2D formulation of Fitts' law, participants thus still had to control all 3D aspects of their hand movements to position the controller correctly.

While ISO 9241-411 [38] standardizes pointing evaluation and throughput calculations, one can question why, as researchers, we should be bound to this standard. The ISO standard was created about 20 years ago and mostly follows Mackenzie's published work [48, 50]. Since then, other researchers have expressed different alternatives. For example, in the standard, W_e is defined as 4.133 SD_x, where SD_x is the standard deviation of movement endpoints in the movement direction. The factor of 4.133 has been shown to be arbitrary by Gori et al. [30]. Morever, Wobbrock et al. [81] showed that bi-variate formulations for *SD* could be used for 2D scenarios. Since the ISO 9241-411 task is (currently) used to assess the user performance for mid-air studies in VR, we invite researchers to further analyze the appropriateness of effective measures to characterize pointing performance in 3D.

Given our results, we suggest that, when using the ISO 9241-411 task to evaluate user performance in 3D user interfaces, authors (at least) mention the task execution strategy they used in their work. Since the effective throughput is not speed-accuracy invariant, the results are affected by the strategy, which might impact the reproducibility of a study. Here, we point out that various 3D pointing studies [19, 39, 57, 67, 76] all used effective throughput in their work but did not share the task execution strategy with the reader. We speculate that the authors of these studies may have considered throughput to be invariant of the task execution strategy. Yet, the findings of our work point out that sharing the task execution strategy is necessary.

5.1.3 Task Execution Strategy and Implications for VR Training

Overall, our results identify that asking participants to focus on different task execution strategies affects their speed, precision, and accuracy. This outcome also indicates that it is possible to use such strategies to affect longer-term training outcomes for users.

In this paper, "active feedback" refers to the experimenter verbally reminding the participants when they were not following the current task execution strategy. This feedback was given more or less immediately, e.g., when participants slowed down too much, they were encouraged to go faster, or when they made too many errors, they were asked to slow down and be more precise. Some motor performance training studies, such as [1, 29, 36, 61, 65] asked participants to execute the task "as fast as possible" i.e., exactly the same instruction as one of the conditions in the experiment. A few studies, such as Sprague et al. [69] or Dresp-Langely [24], asked participants to execute the task "as precise as" possible. A larger set of studies [23, 37, 43, 58, 59] asked participants to execute the task "as fast and as precise as" possible. Based on the fact that all this other work used only these three task execution strategies, we relied on data for these three conditions and observed clear performance differences with these strategies, which also posits sufficient evidence to contradict the findings of MacKenzie and Isokoski [50].

When MacKenzie and Isokoski [50] identified constant throughput, they also mentioned that participants were less precise when they got faster. However, they based this insight only on a single experimental session. Yet, learning is a complex cognitive process that is still not fully understood. For instance, previous work on the speed-accuracy trade-off showed that cognitive tasks in Fitts' law studies can increase the mental fatigue when participants are asked to preserve task success rates independently of the index of difficulty as the task duration increases [64]. The implications of this is that the design of a VR training system can alter the motor performance of trainee. Throughput invariance across varying task execution strategies would lead to constant throughput results. Even thought this seems irrelevant for specific tasks that require either a focus on accuracy, e.g., surgery, or speed, e.g., sports, Batmaz et al. showed that a trainee first must focus on the precision rather than the speed to learn a motor task well [7]. In this case, effective throughput can *only* be increased with training that focuses on precision. While such a claim still require further investigation, we speculate that further research on effective throughput with different users in different application areas can help us to understand how we can use throughput as an assessment criterion for VR training systems.

When analyzing the interaction between ray casting and the virtual hand technique, we saw no difference for the "as precise as possible" task execution strategy in terms of accuracy. Further analysis showed that the effective distance is not significantly different for the ray casting and virtual hand conditions. Even though this was an expected outcome based on previous work [31, 33, 45, 50, 56], our work thus confirms that effective distance has a negligible effect on effective throughput. Specifically, participants' accuracy in the ray casting method reached their accuracy with the virtual hand method when subjects executed the task "as precise as possible". However, this effect is not directly visible in the overall throughput results, since the subjects were overall faster in the virtual hand condition. Yet, we do not know if these results are an outcome of the interaction techniques or the target distances used in our study. Participants selected targets 1.5 m away from their head position with ray casting, which is beyond arm's reach. Thus, these results require further research to reveal whether the target distance or the technique affects results more.

5.2 Limitations

Our work here focuses on how different task execution strategies affect user throughput performance in mid-air pointing. While we investigated the three main task execution strategies used in most previous work, we acknowledge that an investigation of more fine-grained, intermediate task strategies [56] might be a good topic for future work.

We did not focus on comparing user performance of the virtual hand and ray casting pointing techniques, since their differences are well-known. Here we point out that related work [18] already looked not only at the selection techniques but also at their interaction.

Participants used the space bar on the keyboard to select targets. This selection method eliminates the impact of the "Heisenberg effect" on the collected data, i.e., addresses a potential experimental confound [17]. However, this interaction method does not always match how users interact with a virtual environment in application scenarios. Thus, future work should verify if and how the "Heisenberg effect" affects the speed-accuracy trade-off.

The experimental data we used investigated a limited range of IDs, between 2.19 and 4.14. This range was previously used in similar mid-air pointing studies, such as [4, 5, 11, 15]. Nevertheless, even though the ID range was limited, the results of this work still identified a user performance difference for different task execution strategies. Yet, MacKenzie and Isokoski's effective throughput study [50] only investigated a single ID, which could be a possible explanation of the difference between their, Olasdottir et al.'s [56], and our results. This topic should be investigated in future work.

Previous studies on the effect of verbal feedback on motor performance showed that such feedback can improve user performance [3, 60]. However, as Puddefoot et al. states "it should not be assumed that the effect of verbal feedback will be a consistent in every situation." [62]. After all, verbal feedback was based (only) on the experimenter's observations and intended to encourage participants to reach their maximum speed or accuracy. As such, it was not intended to train participants. Also, the number of times active verbal feedback was given during the experiment was not recorded, which might be necessary to investigate VR training methodologies. The impact of verbal feedback on motor performance of participants in an ISO 9241-411 task should thus be studied further.

In this paper, we used data from a study that used a local university participant pool. The average age of the participants was 29.31 $\pm$4.29 years. With other participant groups, such as an older population, it is possible that the outcomes may vary.

6 Conclusions and Future Work

In this work, we examined mid-air VR pointing performance with three task execution strategies with the ray casting and virtual hand selection techniques. Results showed that in contrast to the work that defined throughput and ISO 9241-411, throughput is dependent on execution strategy. This limits its applicability and thus throughput needs to be further investigated before it can be used as a motor performance assessment criterion in situations where the speed-accuracy trade-off plays a key role, such prioritizing training strategies [7] or individual analysis [82]. We hope that our results are helpful for practitioners, developers, and researchers when they aim to assess user performance of their interaction techniques in their VR systems.

In the future, we plan to examine the application of throughput as an assessment criterion for VR training systems and simulators. We are also planning to follow up on the effects of the "Heisenberg effect" on the speed-accuracy trade-off. Moreover, we want to investigate if it is possible to use throughput as a learning assessment criterion and plan to evaluate its effect on the skill transfer from VR to real-world applications.

REFERENCES

[1] G. L. Almeida, D. M. Corcos, and M. L. Latash. Practice and transfer effects during fast single-joint elbow movements in individuals with down syndrome. *Physical Therapy*, 74(11):1000–1012, 1994.

[2] F. Argelaguet and C. Andujar. A survey of 3D object selection techniques for virtual environments. *Computers & Graphics*, 37(3):121–136, 2013.

[3] C. K. Argus, N. D. Gill, J. W. Keogh, and W. G. Hopkins. Acute effects of verbal feedback on upper-body performance in elite athletes. *The Journal of Strength & Conditioning Research*, 25(12):3282–3287, 2011.

[4] M. D. Barrera Machuca and W. Stuerzlinger. Do stereo display deficiencies affect 3D pointing? In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, p. LBW126. ACM, 2018.

[5] M. D. Barrera Machuca and W. Stuerzlinger. The effect of stereo display deficiencies on virtual hand pointing,. In *2019 CHI Conference on Human Factors in Computing Systems*, p. 14 Pages. ACM, 2019.

[6] A. U. Batmaz, M. D. Barrera Machuca, J. Sun, and W. Stuerzlinger. The effect of the vergence-accommodation conflict on virtual hand pointing in immersive displays. In *Conference on Human Factors in Computing Systems*, CHI '22, May 2022. To appear.

[7] A. U. Batmaz, M. de Mathelin, and B. Dresp-Langley. Getting nowhere fast: trade-off between speed and precision in training to execute image-guided hand-tool movements. *BMC psychology*, 4(1):55, 2016.

[8] A. U. Batmaz, M. de Mathelin, and B. Dresp-Langley. Different sound pitch effects on motor performance of individuals in head-mounted virtual reality. In *Perception*, vol. 48, pp. 21–21, 2019.

[9] 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 *Conference on Virtual Reality and 3D User Interfaces*, VR '20, pp. 184–193. IEEE, Mar 2020. doi: 10.1109/VR46266.2020.1581205539914

[10] A. U. Batmaz, A. Mutasim K., and W. Stuerzlinger. Precision vs. power grip: A comparison of pen grip styles for selection in virtual reality, Mar 2020. doi: 10.1109/VRW50115.2020.00012

[11] A. U. Batmaz, M. Rajabi Seraji, J. Kneifel, and W. Stuerzlinger. No jitter please: Effects of rotational and positional jitter on 3d mid-air interaction. In *Future Technologies Conference*, vol. AISC 1289 of *FTC '20*. Springer, Nov 2020. doi: 10.1007/978-3-030-63089-8_52

[12] A. U. Batmaz and W. Stuerzlinger. The effect of rotational jitter on 3d pointing tasks. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '19. ACM, May 2019.

[13] A. U. Batmaz and W. Stuerzlinger. Effect of fixed and infinite ray length on distal 3d pointing in virtual reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–10, 2020.

[14] A. U. Batmaz and W. Stuerzlinger. The effect of pitch in auditory error feedback for fitts' tasks in virtual reality training systems. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 85–94, 2021. doi: 10.1109/VR50410.2021.00029

[15] A. U. Batmaz, X. Sun, D. Taskiran, and W. Stuerzlinger. Hitting the Wall: Mid-air interaction for eye-hand coordination. In *25th ACM Symposium on Virtual Reality Software and Technology*, VRST '19, pp. 30:1–30:5. ACM, New York, NY, USA, 2019.

[16] A. U. Batmaz, S. Xintian, D. Taskiran, and W. Stuerzlinger. Eye-hand coordination training for sports with mid-air vr. In *26th Symposium on Virtual Reality Software and Technology*, VRST '20. ACM, Nov 2020.

[17] D. Bowman, C. Wingrave, J. Campbell, and V. Q. Ly. Using pinch gloves (tm) for both natural and abstract interaction techniques in virtual environments, 2001.

[18] D. Brickler, R. J. Teather, A. T. Duchowski, and S. V. Babu. A Fitts' Law evaluation of visuo-haptic fidelity and sensory mismatch on user performance in a near-field disc transfer task in virtual reality. *ACM Transactions on Applied Perception (TAP)*, 17(4):1–20, 2020.

[19] M. A. Brown and W. Stuerzlinger. Exploring the throughput potential of in-air pointing. In *International Conference on Human-Computer Interaction*, pp. 13–24. Springer, 2016.

[20] Y. Cha and R. Myung. Extended fitts' law for 3D pointing tasks using 3d target arrangements. *International Journal of Industrial Ergonomics*, 43(4):350–355, 2013. doi: 10.1016/j.ergon.2013.05.005

[21] L. D. Clark, A. B. Bhagat, and S. L. Riggs. Extending Fitts' law in three-dimensional virtual environments with current low-cost virtual reality technology. *International Journal of Human-Computer Studies*, 139:102413, 2020. doi: 10.1016/j.ijhcs.2020.102413

[22] S. A. A. Didehkhorshid and R. J. Teather. Selection performance using a scaled virtual stylus cursor in VR. In *Proceedings of Graphics Interface 2020*, GI 2020, pp. 148 – 157, 2020. doi: 10.20380/GI2020.16

[23] S. Doose-Grünefeld, S. B. Eickhoff, and V. I. Müller. Audiovisual emotional processing and neurocognitive functioning in patients with depression. *Frontiers in integrative neuroscience*, 9:3, 2015.

[24] B. Dresp-Langley. Towards expert-based speed–precision control in early simulator training for novice surgeons. *Information*, 9(12):316, 2018.

[25] P. Duc-Minh and W. Stuerzlinger. Is the pen mightier than the controller? a comparison of input devices for selection in virtual and augmented reality. In *25th Symposium on Virtual Reality Software and Technology*, VRST '19. ACM, Nov 2019. doi: 10.1145/3359996.3364264

[26] B. Ens, D. Ahlström, and P. Irani. Moving ahead with peephole pointing: Modelling object selection with head-worn display field of view limitations. In *Proceedings of the 2016 Symposium on Spatial User Interaction*, SUI '16, p. 107–110, 2016. doi: 10.1145/2983310.2985756

[27] I. M. Faniel and T. E. Jacobsen. Reusing scientific data: How earthquake engineering researchers assess the reusability of colleagues' data. *Computer Supported Cooperative Work (CSCW)*, 19(3):355–375, 2010.

[28] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, 47(6):381, 1954.

[29] R. Gentili, C. Papaxanthis, and T. Pozzo. Improvement and generalization of arm motor performance through motor imagery practice. *Neuroscience*, 137(3):761–772, 2006.

[30] J. Gori and O. Rioul. A feedback information-theoretic transmission scheme (fitts) for modeling trajectory variability in aimed movements. *Biological Cybernetics*, 114(6):621–641, 2020.

[31] J. Gori, O. Rioul, and Y. Guiard. Speed-accuracy tradeoff: A formal information-theoretic transmission scheme (FITTS). *ACM Trans. Comput.-Hum. Interact.*, 25(5), Sept. 2018. doi: 10.1145/3231595

[32] J. Gori, O. Rioul, Y. Guiard, and M. Beaudouin-Lafon. The perils of confounding factors: How Fitts' law experiments can lead to false conclusions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, p. 1–10. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3173770

[33] Y. Guiard, H. B. Olafsdottir, and S. T. Perrault. Fitt's law as an explicit time/error trade-off. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, p. 1619–1628. Association for Computing Machinery, New York, NY, USA, 2011.

[34] J. F. Hair Jr, W. C. Black, B. J. Babin, and R. E. Anderson. Multivariate data analysis, 2014.

[35] M. Hartswood, R. Procter, P. Taylor, L. Blot, S. Anderson, M. Rouncefield, and R. Slack. Problems of data mobility and reuse in the provision of computer-based training for screening mammography. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 909–918, 2012.

[36] T. Hazell, K. Kenno, and J. Jakobi. Functional benefit of power training for older adults. *Journal of Aging and Physical Activity*, 15(3):349–359, 2007.

[37] K.-F. Heise, M. Niehoff, J.-F. Feldheim, G. Liuzzi, C. Gerloff, and F. C. Hummel. Differential behavioral and physiological effects of anodal transcranial direct current stimulation in healthy adults of younger and older age. *Frontiers in aging neuroscience*, 6:146, 2014.

[38] ISO 9241-411:2012. Ergonomics of human-system interaction – part 411: Evaluation methods for the design of physical input devices. ISO, May 2012.

[39] S. Jang, W. Stuerzlinger, S. Ambike, and K. Ramani. Modeling cumulative arm fatigue in mid-air interaction based on perceived exertion and kinetics of arm motion. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 3328–3339, 2017.

[40] A. K. Mutasim, W. Stuerzlinger, and A. U. Batmaz. Gaze tracking for eye-hand coordination training systems in virtual reality. Extended Abstract,

Apr 2020. doi: 10.1145/3334480.3382924
[41] B. H. Kantowitz and R. D. Sorkin. *Human factors: Understanding people-system relationships*. Wiley, 1983.
[42] R. Kopper, D. A. Bowman, M. G. Silva, and R. P. McMahan. A human motor behavior model for distal pointing tasks. *International journal of human-computer studies*, 68(10):603–615, 2010.
[43] J. D. Kropotov, V. A. Ponomarev, S. Hollup, and A. Mueller. Dissociating action inhibition, conflict monitoring and sensory mismatch into independent components of event related potentials in go/nogo task. *NeuroImage*, 57(2):565–575, 2011.
[44] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017.
[45] C. N. Liu, C. J. Lin, and C. J. Chao. Fitts' law model and target size of pointing devices in a vibration environment. *Perceptual and Motor Skills*, 105(3):959–967, 2007.
[46] C. L. MacKenzie, R. Marteniuk, C. Dugas, D. Liske, and B. Eickmeier. Three-dimensional movement trajectories in fitts' task: Implications for control. *The Quarterly Journal of Experimental Psychology*, 39(4):629–647, 1987.
[47] I. S. MacKenzie. A note on the information-theoretic basis for fitts' law. *Journal of motor behavior*, 21(3):323–330, 1989.
[48] I. S. MacKenzie. Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction*, 7(1):91–139, 1992.
[49] I. S. MacKenzie. Fitts' throughput and the remarkable case of touch-based target selection. In *International conference on human-computer interaction*, pp. 238–249. Springer, 2015.
[50] I. S. MacKenzie and P. Isokoski. Fitts' throughput and the speed-accuracy tradeoff. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, p. 1633–1636. Association for Computing Machinery, New York, NY, USA, 2008. doi: 10.1145/1357054.1357308
[51] I. S. MacKenzie and A. Oniszczak. A comparison of three selection techniques for touchpads. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 336–343. ACM Press/Addison-Wesley Publishing Co., 1998.
[52] P. Mallery and D. George. SPSS for windows step by step: a simple guide and reference. *Allyn, Bacon, Boston,*, 2003.
[53] P. E. Meehl. Theory-testing in psychology and physics: A methodological paradox. *Philosophy of science*, 34(2):103–115, 1967.
[54] A. Murata and H. Iwase. Extending Fitts' law to a three-dimensional pointing task. *Human Movement Science*, 20(6):791–805, 2001.
[55] J.-Y. Oh and W. Stuerzlinger. Laser pointers as collaborative pointing devices. In *Graphics Interface*, vol. 2002, pp. 141–149, 2002.
[56] H. B. Olafsdottir, Y. Guiard, O. Rioul, and S. T. Perrault. A new test of throughput invariance in Fitts' law: Role of the intercept and of Jensen's inequality. In *Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers*, BCS-HCI '12, p. 119–126. Swindon, GBR, 2012.
[57] M. Pfeiffer and W. Stuerzlinger. 3d virtual hand pointing with ems and vibration feedback. In *2015 ieee symposium on 3d user interfaces (3dui)*, pp. 117–120. IEEE, 2015.
[58] H. P. Pham, R. Chellali, M. Ammi, and J.-g. Fontaine. Frequency modulation based vibrotactile feedback for 1-d and 2-d pointing tasks in virtual reality. In *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 1099–1105. IEEE.
[59] M. D. Popović, M. D. Kostić, S. Z. Rodić, and L. M. Konstantinović. Feedback-mediated upper extremities exercise: increasing patient motivation in poststroke rehabilitation. *BioMed research international*, 2014.
[60] M. C. Porte, G. Xeroulis, R. K. Reznick, and A. Dubrowski. Verbal feedback from an expert is more effective than self-accessed feedback about motion efficiency in learning new surgical skills. *The American journal of surgery*, 193(1):105–110, 2007.
[61] E. Prytz, M. Montano, and M. W. Scerbo. Using fitts' law for a 3d pointing task on a 2d display: effects of depth and vantage point. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 56, pp. 1391–1395. Sage Publications Sage CA: Los Angeles, CA, 2012.
[62] T. Puddefoot, H. Hilliard, and M. Burl. Effect of verbal feedback on the physical performance of children. *Physiotherapy*, 83(2):76–81, 1997.
[63] A. Ramcharitar and R. J. Teather. EZCursorVR: 2D selection with virtual reality head-mounted displays. In *Proceedings of the 44th Graphics Interface Conference*, pp. 123–130. Canadian Human-Computer Communications Society, 2018.
[64] V. Rozand, F. Lebon, C. Papaxanthis, and R. Lepers. Effect of mental fatigue on speed–accuracy trade-off. *Neuroscience*, 297:219–230, 2015.
[65] C. Ruffino, J. Bourrelier, C. Papaxanthis, F. Mourey, and F. Lebon. The use of motor imagery training to retain the performance improvement following physical practice in the elderly. *Experimental brain research*, 237(6):1375–1382, 2019.
[66] F. Sasangohar, I. S. MacKenzie, and S. D. Scott. Evaluation of mouse and touch input for a tabletop display using fitts' reciprocal tapping task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 53, pp. 839–843. SAGE Publications Sage CA: Los Angeles, CA, 2009.
[67] J. N. Shelton and G. T. Chiu. Computing instantaneous throughput for spatially constrained aiming tasks. *IFAC Proceedings Volumes*, 47(3):7208–7213, 2014.
[68] R. W. Soukoreff and I. S. MacKenzie. Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in hci. *Int. J. Hum.-Comput. Stud.*, 61(6):751–789, Dec. 2004.
[69] D. W. Sprague, B. A. Po, and K. S. Booth. The importance of accurate vr head registration on skilled motor performance. In *Proceedings of Graphics interface 2006*, pp. 131–137. Citeseer, 2006.
[70] A. Steed, F. R. Ortega, A. S. Williams, E. Kruijff, W. Stuerzlinger, A. U. Batmaz, A. S. Won, E. S. Rosenberg, A. L. Simeone, and A. Hayes. Evaluating immersive experiences during Covid-19 and beyond. *Interactions*, 27(4):62–67, July 2020. doi: 10.1145/3406098
[71] M. F. Stoelen and D. L. Akin. Assessment of Fitts' law for quantifying combined rotational and translational movements. *Human Factors*, 52(1):63–77, 2010.
[72] R. J. Teather, D. Natapov, and M. Jenkin. Evaluating haptic feedback in virtual environments using iso 9241–9. In *2010 IEEE Virtual Reality Conference (VR)*, pp. 307–308. IEEE, 2010.
[73] R. J. Teather, A. Pavlovych, W. Stuerzlinger, and I. S. MacKenzie. Effects of tracking technology, latency, and spatial jitter on object movement. In *3D User Interfaces, 2009. 3DUI 2009. IEEE Symposium on*, pp. 43–50. IEEE, 2009.
[74] 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)*, pp. 87–94, March 2011.
[75] R. J. Teather and W. Stuerzlinger. Visual aids in 3D point selection experiments. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*, pp. 127–136. ACM, 2014.
[76] R. J. Teather, W. Stuerzlinger, and A. Pavlovych. Fishtank fitts: a desktop vr testbed for evaluating 3d pointing techniques. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, pp. 519–522. 2014.
[77] N. C. M. van Zon, C. Borst, D. M. Pool, and M. M. van Paassen. Touchscreens for aircraft navigation tasks: Comparing accuracy and throughput of three flight deck interfaces using fitts' law. *Human Factors*, 62(6):897–908, 2020. PMID: 31408361. doi: 10.1177/0018720819862146
[78] J. C. Wallis, E. Rolando, and C. L. Borgman. If we share data, will anyone use them? data sharing and reuse in the long tail of science and technology. *PloS one*, 8(7):e67332, 2013.
[79] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, et al. The fair guiding principles for scientific data management and stewardship. *Scientific data*, 3(1):1–9, 2016.
[80] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 143–146. ACM, New York, NY, USA, 2011. doi: 10.1145/1978942.1978963
[81] J. O. Wobbrock, K. Shinohara, and A. Jansen. 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*, pp. 1639–1648, 2011.
[82] S. Zhai, J. Kong, and X. Ren. Speed–accuracy tradeoff in Fitts' law tasks—on the equivalency of actual and nominal pointing precision. *International journal of human-computer studies*, 61(6):823–856, 2004.