A Systematic Review of Fitts' Law in 3D Extended Reality

Mohammadreza Amini
Concordia University
Montreal, QC, Canada
mohammadreza.amini@concordia.ca

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

Robert J Teather
Carleton University
Ottawa, ON, Canada
rob.teather@carleton.ca

Anil Ufuk Batmaz
Concordia University
Montreal, QC, Canada
ufuk.batmaz@concordia.ca

Abstract

Fitts' law is widely used as an evaluation tool for pointing or selection tasks, evolving into diverse applications, including 3D extended reality (XR) environments like virtual, augmented, and mixed reality. Despite standards like ISO 9241:411, the application of Fitts' law varies significantly across studies, complicating comparisons and undermining the reliability of findings in 3D XR research. To address this, we conducted a systematic review of 119 publications, focusing on 122 studies that used Fitts' law in 3D XR user experiments. Our analysis shows that over half of these studies referenced Fitts' law without thoroughly investigating throughput, movement time, or error rate. We performed an in-depth meta-analysis to examine how Fitts' law is incorporated into research. By highlighting trends and inconsistencies, and making recommendations this review aims to guide researchers in designing and performing more effective and consistent Fitts-based studies in 3D XR, enhancing the quality and impact of future research.

CCS Concepts

• **Human-centered computing → Human computer interaction (HCI)**; **HCI design and evaluation methods**; **User models**; **HCI theory, concepts and models**; **Mixed / augmented reality**; **Virtual reality**.

Keywords

Fitts' Law, Systematic Review, Extended Reality (XR), Virtual Reality (VR), Augmented Reality (AR), Mixed Reality (MR)

ACM Reference Format:
Mohammadreza Amini, Wolfgang Stuerzlinger, Robert J Teather, and Anil Ufuk Batmaz. 2025. A Systematic Review of Fitts' Law in 3D Extended Reality. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26-May 1, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 25 pages. https://doi.org/10.1145/3706598.3713623

CHI '25, April 26-May 1, 2025, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713623

1 Introduction

Fitts' law, introduced by Paul Fitts in 1954, provides a predictive model for rapid aimed movements [52], specifically describing the relationship between Movement Time (MT) and the difficulty of a target acquisition task. This model originated in human motor control studies, where it described how MT increases with both larger target distance and smaller target size. Over the years, Fitts' law has gained significant traction in Human-Computer Interaction (HCI) research due to its usefulness as an evaluation tool, and remarkable ability to reliably model pointing interactions [94].

In HCI, Fitts' law was initially applied to 2D graphical user interfaces to evaluate tasks like pointing, clicking, and dragging, guiding the development of interaction techniques and devices such as mice, touchpads, and styli [35, 97]. Over time, it has become a key framework for evaluating various devices and interfaces, from desktop computers to touchscreens, and for designing interaction methods that improve user performance [1].

With steadily increasing research focus on Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR) systems, collectively referred to as Extended Reality (XR) systems, new aspects of HCI have emerged. XR systems are increasingly used in various aspects of people's lives, from gaming and education to healthcare and industrial applications [6, 69, 127, 147]. As XR technologies continue to evolve, XR interaction design becomes increasingly more complex, making Fitts' law a valuable tool for assessing user performance in these new environments [56]. Fitts' law also serves as an effective tool for designing interaction systems in this context, and, as XR research advances, new extensions of the law are being explored to accommodate the complex characteristics of XR, expanding its applications and effectiveness.

While Fitts' law was initially applied to 1D (one-dimensional) pointing tasks, its application expanded to 2D, such as for the mouse and touch screens, and into 3D, such as for virtual hand and interaction spaces, which are common in XR environments [30, 143]. This expansion has led to a wide range of variations, with adaptations and extensions of Fitts' law developed to address the unique complexities of XR interactions [134, 145]. As a result, a variety of methodologies, experimental designs, and evaluation metrics inspired by Fitts' law have emerged for use in XR systems.

However, this diversity also presents challenges for 3D User Interface research [81, 145]. While Fitts' law serves as a common framework for evaluating and comparing interaction techniques, the numerous variations and adaptations for different studies make it difficult—if not impossible—to directly compare results [53, 71].

The lack of consistency in experimental setups, measurements, and evaluation methods leads to a range of discrepancies between results, which in turn makes it challenging to draw meaningful conclusions for practical applications, as the results vary too much to be able to tell how the results of a specific study match other work. This creates a need for a more cohesive understanding of how Fitts' law is applied in XR and HCI research and what trends, gaps, and inconsistencies exist.

This paper aims to identify trends and patterns in Fitts' law research within 3D XR, focusing on commonalities and variations in user studies. We seek to assess the consistency among existing studies and highlight sources of variability while identifying gaps and opportunities for further research. We aim to advance the understanding of Fitts' law and its applications in 3D XR. We also provide insights based on the current state of literature through an actionable research framework, ultimately improving the quality of research and moving towards standardization of future studies.

2 Related Work

Fitts' law is a predictive model of rapid aimed movements and is widely used in HCI for different purposes. Two main usages are designing user interfaces [27] and evaluating new interaction methods [94]. Originally, Fitts proposed that there is a linear relationship between movement time and the ratio of target width (W) to distance (typically A, for amplitude) in the reciprocal tapping task (where a and b are linear regression coefficients):

$$MT = a + b\log_2\left(\frac{2A}{W}\right) \tag{1}$$

The logarithmic term is referred to as the Index of Difficulty (ID), which indicates the overall difficulty of the task based on the distance to and size of the target (Equation 2 [52]). The Index of Performance (IP), more commonly now referred to as throughput (TP), is the information transmission rate and the primary tool of Fitts' law to compare the performance of various input techniques or devices [95] when evaluating a series of target selection tasks. TP was originally proposed as the inverse of the slope (1/b) but to avoid confusion with the more modern definition of TP (Equation 3 [94]), we suggest calling the reciprocal of the slope consistently IP.

$$ID = \log_2\left(\frac{2A}{W}\right) \quad [bits] \tag{2}$$

$$TP = ID/MT \quad [bits/s] \tag{3}$$

Since then, there have been many refinements to the original formulation. Welford's [154] formulations (Equation 4), and Equation 5 by Meyer et al. [103] were the earliest variations of the law, with the so-called Shannon formulation proposed by MacKenzie [94] as likely the most commonly used today (Equation 6). Each variant was devised to improve the correlation and accuracy of the model by adjusting the ID.

$$MT = a + b\log_2\left(\frac{A}{W} + 0.5\right) \tag{4}$$

$$MT = a + b\sqrt{\frac{A}{W}} \tag{5}$$

$$MT = a + b\log_2\left(\frac{A}{W} + 1\right) \tag{6}$$

As extensions to Fitts' law grew with the introduction of new models, the concept of effective width (Equation 7 [41, 94]), became a part of ISO 9241-9 [66] and later the ISO 9241-411 standards [67]. As it adjusts the Fitts' law based on the selection noise [95] to improve the accuracy aspect in the speed-accuracy tradeoff, this refined version of the model gained substantial traction in HCI research.

$$W_e = 4.133 \cdot SD_x \tag{7}$$

In Equation 7, SD_x is calculated as the standard deviation between the target center and a set of selection positions for the same task, giving the "spread" of selection coordinate distances along the task-axis, i.e., the line between subsequent targets [16, 98, 99]. Based on the ISO 9241:411 standard, this is multiplied by 4.133, which corresponds to ±2.066 standard deviations of a normal distribution, which accounts for 96% of the values under the "bell" curve [94]. This effective width measure (W_e) post-experimentally adjusts the size of targets such that 96% of the selections would have hit the target - corresponding to an error rate of 4% [94]. This adjustment facilitates the comparison of study outcomes with otherwise varying error rates. The ISO standards [66, 67] recommend the use of effective width in deriving an effective index of difficulty (ID_e), and correspondingly using this to calculate throughput as seen in Equation 8 [94] and Equation 9 [94].

$$ID_e = \log_2\left(\frac{A_e}{W_e} + 1\right) \tag{8}$$

$$TP = \frac{ID_e}{MT} \tag{9}$$

Although there were attempts to revisit and standardize Fitts' law in HCI studies (e.g., [95, 100, 132]), not having a single method or standards for evaluating more complex interactions remains controversial. Notably, in 3D XR, even a "simple" Fitts' law study is complicated by factors such as the definition of the (visual) depth of the target placement (i.e., how far the targets are from the viewer), the presence or absence of stereo viewing, unclear definitions of what constitutes the selection coordinate (e.g., in consideration of ray vs. hand-based selection, where rays usually consider the object/ray intersection point as the selection coordinate, while virtual hands usually use the actual 3D position of the hand/controller as the selection coordinate), and various and non-standard selection methods. There are numerous extensions (e.g., [12, 40, 108]), that introduce new parts to the original model, focusing on 3D interaction, such as Equation 10 by Machuca and Stuerzlinger, who introduced *CTD* to take the combined effect of distance and size in 3D into account [12].

$$MT = a + b\,(ID) + c\,(CTD) \tag{10}$$

Two of the most widely used and standard interaction methods in XR are virtual hand and ray casting [81, 116, 135, 149]. Virtual hand selection is more similar to the original task that Fitts used since it directly maps the "cursor" position to that of the hand. Nevertheless, there are many variations of Fitts' law for near-field and distal target pointing and selection, and no single standard formulation exists [12, 16, 37, 40, 108]. Furthermore, with other interaction methods like ray casting, new characteristics are added to Fitts' original selection. These interactions usually extend from the input position and do not directly map the position of the hand. For the adaption of Fitts' law to these kinds of interactions, new models have emerged, such as Kopper et al.'s variant [78] that extend the original law to rotational movements (Equation 11), making such versions a common choice for research on ray casting (or similar input) methods, especially those incorporating angular distances [16, 78].

$$MT = a + b \log_2 \left(\frac{\alpha}{\omega^k} + 1 \right) \tag{11}$$

In Equation 11, α and ω are the angular target distance and angular target width, respectively. The k term indicates the relative weights of the target's angular distance and width. In addition, the definition of effective width is also applied in this equation for the calculation of effective TP with angular measures, such as in Batmaz et al.'s work [16].

Teather and Stuerzlinger's paper, "Pointing at 3D Targets in a Stereo Head-Tracked Virtual Environment" [143] is one of the earliest applications of Fitts' law in 3D XR, incorporating ISO 9241-9 in their user study. Since then, numerous studies have employed similar methodologies to investigate pointing and selection in 3D environments. Among the first reviews of selection techniques is Hand's survey of 3D interaction techniques [58]. He focused on different interaction techniques designed for mouse and 3D input devices. When mentioning Fitts' law, Hand stated that "***the HCI community has little in the way of standard evaluation methods for 3D interaction short of adopting techniques such as Fitts' law***". A later study by Bowman et al. [31] classified 3D selection techniques by different tasks. However, they did not focus on or investigate Fitts' law in their literature review. Dang also surveyed and classified 3D selection techniques and how 2D selection techniques applied to 3D user interfaces in 2007 [43].

In recent years, surveys on 3D interaction techniques mostly focused on immersive environments and VR/AR/XR HMDs. In Arelaguet and Andujar's survey [9], along with 3D object selection method classification and human models used for evaluation, the authors examined the factors influencing user performance. They investigated Fitts' law as a methodology to model human movement in 3D environments but did not investigate the details in practice. Triantafyllidis and Li [145] reviewed extensions of Fitts' law highlighting the application complexities in 3D environments, especially for combined rotational and translational tasks, providing an understanding of the limitations of current 3D Fitts' law extensions. In Mendes et al.'s survey [102], the authors analyzed different trends in interaction and object manipulation in virtual environments but did not investigate Fitts' law. Similarly, other papers surveyed 3D interaction techniques, such as Subramanian and IJsselsteijn [140] and Zhi et al. [162], but these papers did not focus on Fitts' law and how it is applied to 3D interaction in their reviews. Here, we focus on recent practices around Fitts' law in 3D XR research.

3 Motivation and Objectives

XR research is rapidly advancing with new technologies and interaction methods. At the same time, Fitts' law has become a crucial component of designing and evaluating such interaction methods through formal experiments. As seen in Figure 1a, the increasing number of publications using Fitts' law in user studies highlights the growing role of Fitts' law in the field. A similar explosion in XR research is also evident (Figure 1b).

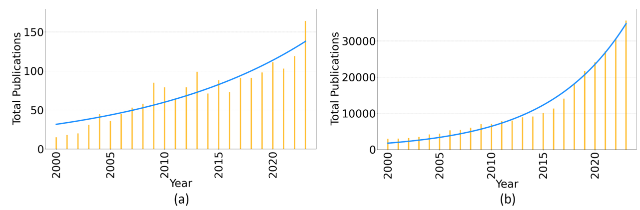

Figure 1: (a) Rise of Fitts' law usage in publications during recent years (2000-2023). (b) The rapid growth of XR (including VR, AR, MR) during this period. (ACM digital library, IEEE Xplore, Elsevier ScienceDirect, and SpringerLink)

Given the increasing use of Fitts' law in XR research, a comprehensive literature review to summarize the many previous studies and provide guidelines for future ones is warranted. In this paper, we present a systematic literature review, aiming to synthesize research methodologies using Fitts' law and the corresponding findings to address the emerging need to limit future inconsistencies within XR Fitts' law studies and point out gaps in the literature. To address this, through an in-depth meta-analysis, we aim to accomplish several key objectives:

- Identify trends and patterns in the rapidly growing set of Fitts' law studies and extensions in the context of 3D XR, identifying common ground.
- Provide a quantitative comparison and analysis of various metrics, including MT, TP, and Error Rate (ER), across different interaction techniques and XR displays.
- Investigate and assess the general consistency among existing studies while highlighting sources of variability.
- Identify gaps and opportunities in the current literature and suggest areas for further investigation and study.
- Provide a framework to conduct higher-quality Fitts' law studies in 3D XR.

4 Methodology

In conducting our systematic literature review, we employed the PRISMA-2020 guidelines [111]. PRISMA-2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) is a set of guidelines designed to improve systematic literature reviews by making their reporting more transparent and easy to understand. Originally developed for medical research, PRISMA has since been widely adopted across various fields, including Human-Computer Interaction (HCI), and has shown to be successful in producing high-quality literature reviews, e.g., those by Agha et al. [4], Leclercq et al. [83], Panic et al. [114], and Stefanidi et al. [136].

Our methodology followed the PRISMA framework to ensure that our results are reliable and reproducible. By following the PRISMA-2020 principles, we systemically uncovered, evaluated, and assessed relevant studies, improving the validity and dependability of our conclusions. The review was conducted in three main phases:

- *Initial capture:* We collected a wide range of papers that at least mentioned Fitts' law.
- *Detailed analysis:* In this phase, we performed an in-depth analysis of all user studies from the initial set, focusing on how Fitts' methodology for selection tasks was applied and what data was recorded by the researchers.
- *General analysis:* Papers not included in the detailed analysis were categorized to identify trends and applications of Fitts' law beyond the mentioned user studies in our context.

4.1 Information Sources

To gather the relevant literature on the applications of Fitts' law within 3D XR research, our search was conducted across four major digital libraries—ACM Digital Library [3], IEEE Xplore [65], SpringerLink [133], and Elsevier ScienceDirect [48], due to their extensive collections of reliable, peer-reviewed studies and their popularity within the field. We did not limit our sources to specific publications or dates, aiming for broad coverage of relevant research, including interdisciplinary and high-quality papers that used Fitts' law in 3D.

4.2 Search Strategy, Selection Process, and Eligibility Criteria

4.2.1 Identification. As illustrated in Figure 2, we began by retrieving papers from multiple databases on July 31st, 2024. We used a query that is broad enough to capture all Fitts' law studies in XR. Here is the list of queries used for finding related publications in each database:

- ACM digital library: `[All: "fitts law"] AND [[All: "virtual reality"] OR [All: "3d"] OR [All: "augmented reality"] OR [All: "mixed reality"] OR [All: "extended reality"]]`.
- Elsevier ScienceDirect: `"fitts law" AND ("virtual reality" OR "augmented reality" OR "mixed reality" OR "extended reality" OR "3d")`
- IEEE Xplore: `(("Full Text & Metadata":"virtual reality") OR ("Full Text & Metadata":"augmented reality") OR ("Full Text & Metadata":"mixed reality") OR ("Full Text & Metadata":3D) OR ("Full Text & Metadata":"extended reality")) AND ("Full Text & Metadata":"fitts' law")`
- SpringerLink: `"fitts law" AND ("virtual reality" OR "augmented reality" OR "mixed reality" OR "extended reality" OR "3d")`

In total, these queries identified 1957 papers. Among these papers, we excluded 491 papers based on language (12 papers) and content type (479 papers), i.e., journal articles, conference papers, and book chapters, to ensure the inclusion of complete and peer-reviewed research. Shorter formats such as abstracts, posters, and workshop papers were excluded due to, e.g., their lack of depth and peer review. Nonarchival materials and other content that could later be expanded into full articles were also excluded, as their results might change. This refined set of papers was then subjected to a manual screening process.

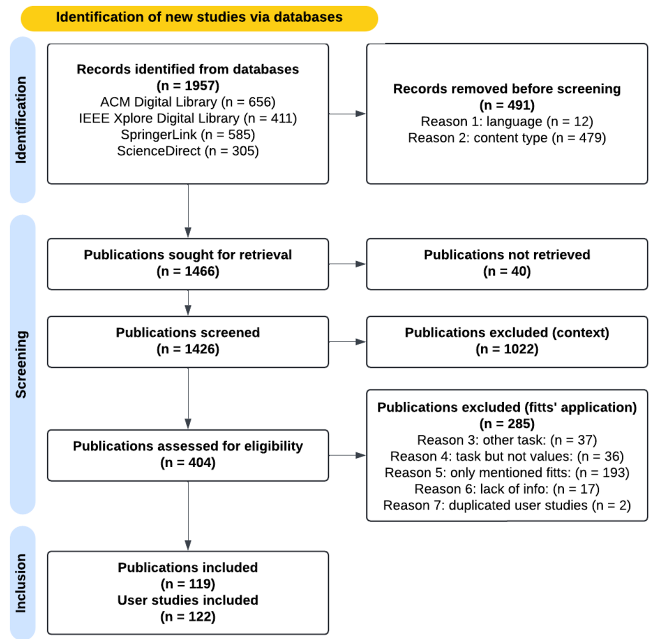

Figure 2: Study selection process flow-diagram used for this study.

4.2.2 Screening. During the screening, we encountered access issues for 40 papers due to licensing limitations. Yet, all these were book chapters, and their titles and abstracts indicated they did not include any user studies conducted by the authors. One author manually screened the full text of the remaining 1426 papers to filter those relevant to our focus on user studies in 3D XR. This author also checked the papers to prevent misinterpretations, and vetted selection tasks or tasks related to selection, like pointing, to ensure comparability and because this task aligns with the original Fitts' task design [52]. The filtering process excluded 1022 papers based on the context that matched the query but were irrelevant to the study, such as papers that used the words "Fitts' law" or "3D" in their metadata, documents that did not contain user studies like general survey papers, or those that were hit by the query but did not incorporate a 3D perspective view.

4.2.3 Inclusion. We also evaluated the remaining 404 reports against our eligibility criteria for inclusion. We excluded reports for five main reasons: (1) they investigated different tasks (e.g., moving targets [36, 64], steering law [88, 160] or peg-and-hole task [8]) or (2) did not use Fitts' law, or applied Fitts' law to study other values except TP, ER, and MT, which are outside the scope of our investigation (e.g., physical attributes such as shoulder angle [77] and task load and discomfort [115]). Other excluded reports in this phase either (3) mentioned Fitts' law without using it as a methodology (e.g., related work by others [92], general claim [84], and future work [26]), (4) lacked sufficient information on the data items and methodology to be effective in our study (e.g., [39, 82]), or (5) exhibited some combination of these above reasons. In addition, full-text screening helped us recognize duplicate publications (e.g., [33, 137]). Ultimately, we included 119 publications in the final in-depth analysis to extract data related to their Fitts' law user studies in XR. When papers presented different user studies with separate experimental designs (e.g., one experiment with ISO 9241-411 multidirectional tapping task [67, 95] and one Fitts' original bar selection task [52]), we extracted values for each study individually, resulting in data for 122 user studies.

4.3 Data Collection Process and Measures

For analyzing Fitts' methodology in user studies, we recorded variables from the papers and aimed to extract the most reliable values. Details about the variables that we collected are provided in Table 1. When data values were visualized in charts, we extracted only a range based on the granularity of the corresponding axis. Due to the varied presentation of experiment-related values, we performed all data extraction manually and peer-reviewed it to ensure the accuracy and quality of the retrieved information. The data collection process occurred in two cycles: first, we captured detailed data to identify trends and develop a framework for different data types and representations; second, we used this framework to merge and format the information for the final synthesis and analysis.

We also observed varying terminologies for data items in Fitts' law studies, such as throughput and index of performance, which often referred to the same measure. During data collection, we included values reported with different names but referring to the same measure. In contrast, we did not report values for measures with similar names that did not match the data items we are investigating. For instance, we included error rate (ER) but excluded measures like accuracy due to the difficulty of reliably converting these values to ER across studies. As for the movement time, we included values for metrics like selection time (if they were referring to the data item under study), but did not capture task total completion time to decrease the variety of reported values and keep data collection consistent.

4.4 Risk of Bias and Certainty Assessment

In our study, one potential source of bias stems from duplications, where the same study is published across different libraries. The large number of papers reviewed presents a challenge for accurately identifying and managing such duplications, by both human reviewers and automated systems. Still, we made efforts to minimize duplications by removing obvious duplicate user studies encountered, based on memory. We also further investigated papers with similar results in the data synthesis phase to minimize bias.

Our focus on user studies involving 3D perspective views in XR, such as head-mounted and stereoscopic displays, helped narrow the scope and reduced biases from irrelevant studies (e.g., those using standard 2D wall projections or laptops without stereoscopic view). This inclusion criterion further ensured that our analysis concentrated on relevant studies within the desired field, thereby mitigating potential biases in our results. Besides, since all reviews in this study were manually done by human reviewers over a substantial period, potential inconsistencies may have been created over time. In addition, some user studies may report less accurate results or be contaminated by biases (such as the bias in non-reported results). Yet, due to the large volume of papers, it was impractical to investigate biases within each study, including identifying flaws in experimental design. As a result, we relied on the reported values by the authors. This is another reason why we focused on high-quality peer-reviewed publications in the identification phase.

4.5 Synthesis Methods

After extracting information from the papers, we merged the data into a format suitable for analysis in Excel. The data for each user study was consolidated into a single row, which was then passed to the analysis phase. For data items where we only focus on the percentage of the results (like ID exact, where we aim to investigate what portion of studies reported the exact values for ID), we provided the partial values (like 0.5, when two experiments are done using same target arrangements, but ID is reported for one of the experiments) in the merged cell to present more precise results. For each study that used different metrics for measures like target size and distance in their experiment design, we first performed an overview to identify the emerging trends and then prioritized the most popular metrics for our cross-study analysis (e.g., if a study used both angular and Euclidean metrics for target distance, we reported the Euclidean value, since it is more common between investigated studies). Similarly, in studies that used or mentioned different equations derived from Fitts' law (e.g., the classic formulation Equation 2 or Shannon's Equation 6), we first identified the formula used for throughput and related calculations, prioritizing the most commonly applied version. This approach ensured the captured values were comparable across a broader range of studies..

5 Results

5.1 Fitts' Variations

Fitts' law and its extensions have been a topic of debate in 3D XR research, as the original equation was developed for 1D tasks and later *adapted* for 2D and 3D interactions. The choice of the appropriate formula is critical, as it can significantly influence a study's outcomes. During our full-text data retrieval and meta-analysis, we thus observed inconsistencies and ambiguities regarding the formulas used. While many studies reference different variations of Fitts' law, they often apply only one version, which may not always be explicitly stated. This highlights the importance of clearly indicating which formula is used to generate results, as it poses a challenge for cross-study comparisons and future research. For example, in Kopper et al.'s formulation (Equation 11) set $k = 1$ to transform it into the Shannon formulation (Equation 6), but using angular measures. Although the k constant in Kopper et al.'s formula has been a topic of discussion in the literature, our focus is on reporting the values used, so we do not explore this issue further. Additionally, Kopper et al.'s approach is commonly used in studies employing the angular formulation to ensure comparability with results based on Shannon's version (e.g., [122]). In Table 2, we present the extensions of Fitts' law identified during our full-text meta-analysis along with their corresponding percentage of occurrences.

Table 1: Summary of data items and their representations in the study

Data Item	Data Type	Example	Details
Throughput (TP) / Movement Time (MT) / Error Rate (ER) Captured	Boolean	1/0	If TP / MT / ER is measured in the user study
Index of difficulty (ID) reported	Boolean	1/0	If ID is reported
TP (bits/s) / MT (s) / ER (%) / ID (bits) Range	Numeric Values	(3.25,3.75)	Range provided in charts, exact average value if specified
TP / MT / ER / ID Exact	Boolean	1/0	If they reported the exact value for it
TP / MT / ER / ID Scale	Numeric Values	0.5	If a diagram is used, what is the most precise step size of the diagram
Number of IDs	Numeric Values	10	Number of IDs used
Feedback Type	Text	visual	Feedback used in the experiment design
Regression Reported	Boolean	1/0	If they reported the Fitts' regression
Slope (b) Range	Numeric Values	0.255	Range provided in charts, exact average value if specified
Environment	Text	VR (HTC Vive)	Experiment environment, (device) if specified
Selection Technique	Text	ray casting	Selection technique used
Condition	Text	feedback evaluation	General aim of the study
Target Amplitude (A) and Width (W) Reported	Boolean	1/0	If A and W used for experiment design is reported
A and W Metric	Text	angular	Main metric used for A and W
Distance From User / Distance From Screen (cm)	Numeric Values	(300,600)	Target (visual) depth information
Experiment Design	Text	grid target selection	Task used for the study
Participant Info	Text	10 (6,4,2) 23.5	Format: number (male,female,other) mean_age

Due to the ambiguities discussed in this paper, the percentages for extensions 3 and 4, as well as 7 and 8, have been merged. Not all of these extensions are necessarily used to evaluate 3D selection tasks; some are mentioned as variations not directly applied in the study but evaluated as models or used as a basis for investigating new models. For example, Clark et al. [40] investigated extensions 9-12 (Table 2) to explore potential factors influencing MT in VR. Janzen et al. [70] explored extensions 5-8 (Table 2) to study the effect of target depth on performance in target-pointing tasks.

Our meta-analysis showed that the effective-measures formulation of Fitts' law is the most mentioned and studied version. It was explicitly referenced in 61 (50.0%) studies, no matter what metric they used for target amplitude and size (e.g., [15, 18, 50, 54, 106, 112, 125, 150]). In addition, 16 other studies did not report the formula they used precisely, but stated that they followed ISO 9241 in their user studies (e.g., [75, 86, 90]). Since the use of effective measures is mentioned in the ISO 9241 document, this fact could imply that these studies used effective calculation too, increasing the number of studies that incorporated this formula to 78 (63.93%). There are other studies, such as Yu et al. [158] that mentioned Shannon's formulation in the text, but did the calculations based on other studies (Soukoreff and MacKenzie[132] in the mentioned example), which likely means they did follow the effective variation. In total, 25 (20.5%) studies were identified that only mentioned Fitts' law or ISO 9241 without directly reporting the variation used in their calculations and/or if they used effective measures. This makes a cross-study evaluation challenging for researchers since there is no indication if the accuracy adjustments have been adapted or not. Nevertheless, the effective-measures formulation is clearly the most widely used variation of Fitts' law.

In contrast, with the advances in 3D XR, the original Fitts' law formulation (Equation 2) is rarely used anymore. Traces of classic formulation can only be seen in older studies (e.g., [152], which used the classic calculation of ID, not the MT equation). Even studies that briefly mentioned the classic formulation without explicitly using it in their methodology (e.g., [85]) are scarce, with usages accounting for less than 5%.

Table 2: Fitts' Law variations in 3D XR studies.

#	Name	Equation	% of Studies
1	Original [52]	$MT = a + b\left(\frac{2A}{W}\right)$	4.0%
2	Meyer et al. [103]	$MT = a + b\sqrt{\frac{A}{W}}$	< 2.0%
3	Shannon [94]	$MT = a + b\log_2\left(\frac{A}{W} + 1\right)$	77.8%
4	Effective [94]	$MT = a + b\log_2\left(\frac{A_e}{W_e} + 1\right)$	
5	Two-part [154]	$MT = a + b_1\log(A) - b_2\log(W)$	< 2.0%
6	Shannon-Welford [129]	$MT = a + b_1\log(A + W) - b_2\log(W)$	< 2.0%
7	Angular [78]	$MT = a + b\log_2\left(\frac{\alpha}{\omega^k} + 1\right)$	28.7%
8	Effective Angular [16]	$MT = a + b\log_2\left(\frac{\alpha_e}{\omega_e^k} + 1\right)$	
9	Hoffmann [60]	$MT = a + b\log_2(\frac{2R}{S+P})$	< 2.0%
10	Murata and Iwase [108]	$MT = a + b(\sin\varphi) + c\log_2\left(\frac{A}{W} + 1\right)$	< 2.0%
11	Cha and Myung [37]	$MT = a + b\varphi + c\sin(\theta) + d\cdot\log_2\left(\frac{2D}{W+F}\right)$	< 2.0%
12	Machuca and Stuerzlinger [12]	$MT = a + b(ID) + c(CTD)$	3.2%
13	MacKenzie and Buxton[97]	$MT_{min} = a + b\log_2\left(\frac{A}{\min(W,\alpha_1 H)} + 1\right)$	< 2.0%
14	Accot and Zhai [2]	$MT = a + b\log\left(\sqrt{\left(\frac{A}{W}\right)^2 + \alpha_1\left(\frac{A}{H}\right)^2} + 1\right)$	< 2.0%

Additionally, for the calculation of the TP, studies mostly used $TP = ID/MT$, rather than the traditional calculation based on the slope ($TP = 1/b$). In our analysis, roughly 3% of studies used only the IP, i.e., the reciprocal of the slope for the calculation of "TP", e.g., Wolf et al. [156].

Challenges encountered during the analysis of different Fitts' law variations in our meta-analysis highlight the need for clear specification of the extensions used. Table 2 presents some of the most cited extensions of Fitts' law in 3D XR studies. These extensions are evolving alongside XR advancements, each addressing specific factors, such as visual depth, to enhance the modeling of user behavior.

5.2 Fitts' Applications

There are various applications and usages for Fitts' law in the HCI context [71, 94]. This diversity is also observed in 3D XR studies throughout the different phases of this systematic review. In our in-depth meta-analysis of the included papers, we categorized the applications of Fitts' law into three distinct aspects (Table 3). The studies analyzed in this phase are those that specifically employed Fitts' law as part of their user studies and applied it to at least one of the identified categories.

As seen in Table 3, the studies analyzed mostly used Fitts' law to evaluate an interaction method or a system. It is also used for evaluating different feedback types such as auditory (e.g., [19, 24]) or haptic feedback (e.g., [17, 141]). Various studies also investigated the effect of other factors on user performance and interactions, such as Batmaz et al. [18] investigating the effect of rotational jitter, Kohli et al. [76] who studied the impact of warped space on task performance, or Machuca and Stuerzlinger [12] studying the effects of stereo display deficiencies.

The use of Fitts' law in 3D XR extends beyond the areas mentioned above. There are numerous other applications of Fitts' law within the context of XR, which may not fall within the scope of our in-depth analysis of user studies incorporating a 3D perspective but are still worth mentioning. To provide a broader view of how Fitts' law has been employed in XR research, we further examined papers that were screened but excluded from our detailed meta-analysis due to incompatible applications of Fitts' law or a lack of a 3D perspective.

5.2.1 Accessibility. With the advances in XR, it is becoming an effective technology with promising results and opportunities for users with different accessibility needs, such as people undergoing rehabilitation [62, 123, 139]. Our systematic review shows that Fitts' law is also applied in research related to accessibility for designing and evaluating systems. For example, Jacho-Guanoluisa et al. [68] applied Fitts' law to a cognitive rehabilitation application designed for children with rare diseases. Fitts' law was also applied in many other papers screened during this systematic review aiming to design and investigate rehabilitation systems (e.g., [11, 49, 61, 105, 113, 165]).

Past research has also shown that Fitts' law is well-suited to evaluating selection tasks for people with visual impairments [74, 89]. For example, Hu et al. [63] used Fitts' law to evaluate their wearable target location system, designed for visually impaired or blind users. Fitts' law is also used in studies for people with motor impairments. Vatavu and Ungurean [148] used a Fitts' evaluation to investigate input performance for people with upper-body motor impairments. In another study, Zimmerli et al. [166] implemented a mechanism based on Fitts' law to balance the difficulty of upper-extremity rehabilitation tasks. These examples illustrate that Fitts' law is

Table 3: Main applications of Fitts' law in investigated user studies.

#	Focus of Study	% of Studies	Examples
1	Interaction and System Evaluation	62.3%	[50, 54, 104, 150, 156, 158, 161]
2	Feedback Evaluation	13.9 %	[17, 19, 24, 125, 126, 141]
3	Other Factors Influence on Interaction	26.2 %	[12, 14, 18, 23, 93, 146]

also useful for researchers focusing on designing and evaluating interaction systems for users with accessibility difficulties.

5.2.2 Medical Education and Training. As XR technology advances, it is increasingly being explored as a powerful and cost-effective tool for medical education in different areas, such as anatomical training, self-paced safe learning, and also surgical and neurosurgical training [42, 121, 130, 167]. In addition, with the increased need for remote education and training during the COVID-19 pandemic, XR technologies such as VR have shown to be effective, especially for medical training [7, 151, 167]. With the growing popularity of XR in education, especially medical education, comes the need for effective design and evaluation of such systems. Throughout our investigation, we identified studies (e.g., [45, 46, 79, 80]) that used Fitts' law as a design and evaluation tool in exploring XR technologies for medical purposes. This finding further highlights the applicability of Fitts' law in this growing context.

5.2.3 Human-Robot Interaction (HRI). During our systematic review, we also identified a large number of papers used Fitts' law for either designing or evaluating Human-Robot Interaction (HRI) methods in 3D XR [28, 47, 72, 79, 119, 120, 128, 166]. Besides, XR is being widely explored as a means for HRI for building innovative interaction methods that have future promise [44, 57, 155]. XR-based HRI is thus another example of Fitts' law's vast area of applications. For instance, Prada et al. [117] evaluated their MR solution for HRI using Fitts' law.

5.3 Environment (Display and Interaction Method)

Our results synthesis reveals that various display systems are used in 3D XR Fitts' studies which is to be expected with the advancing nature of this field. In total, 83.5% of papers explored HMDs in their studies (e.g., [15, 17, 91], which are mostly either VR (79.4%) and AR or MR (20.6%) systems. The rest, 19.7% of the papers, incorporated more traditional stereo-displays rather than HMDs (e.g., [70, 142, 143]).

Figure 3 illustrates the disparity and trend of reported values for MT, ER, and TP based on the display explored. In addition, the scatter plot inside each category represents the number of studies and their distribution. The median TP values are 3.25 bits/s and 3.20 bits/s for HMD and stereo-displays respectively. As for the MT, median values are 1.25s (HMD) and 1.32s (stereo-display), and for the ER 11.45% (HMD) and 3.20% (stereo-display), respectively. Also, in Figure 3 stereo-displays exhibit significantly less disparity for the recorded ERs in Fitts' task studies.

In addition to different displays used in XR Fitts' studies, numerous interaction methods are explored, acting as another source of variability. Of the studies reviewed, 36% used the virtual hand

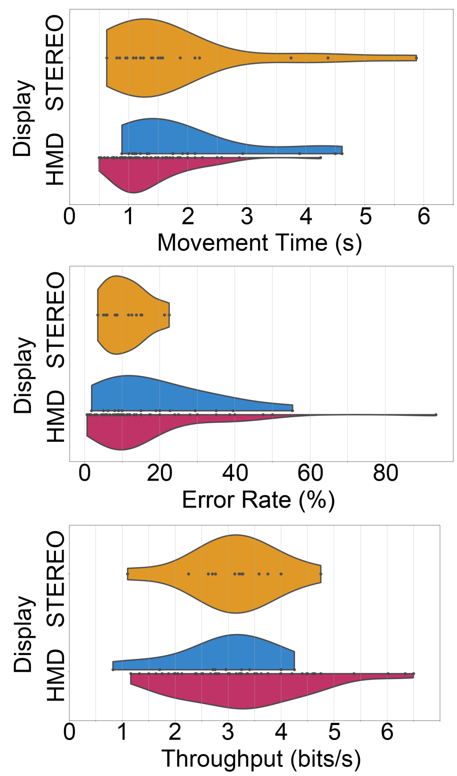

Figure 3: Trend of reported values for Movement Time (MT), Error Rate (ER), and Throughput (TP) based on the display type, i.e., Head-Mounted Display (HMD) or Stereo-display (Stereo). Distinct colors indicate AR (blue) and VR (Red) systems.

(e.g., [15, 17]) and 28.6% used ray casting (e.g., [18, 146]) as a part of their interaction method. Each different interaction (selection) method involves different characteristics in aimed selection, like the distance to the target, resulting in variance in methodologies.

5.4 Measures

5.4.1 MT, ER, and TP. Our meta-analysis found that MT is the most common metric reported in 3D XR Fitts' user studies. For each metric, we assessed how precise values are reported in publications since a reliable and informative perception of reports is critical. In Table 4, besides the number of studies that measured each metric among investigated studies, we list how many reported the exact value (either in diagrams or in text), and in those that incorporated diagrams for showing the results, what the average step size of the corresponding axis is. Furthermore, the same inconsistency in studies is also observed in reporting ID values (71% reported ID range, 49% reported exact ID values, and the median step size in diagrams with IDs as an axis is 0.5 bits). For a better representation of popularity, Figure 4 illustrates the precision trend in reports for MT, ER, and TP. It shows that with a maximum step size of 0.5 s, 5%, and 0.5 bits/s for MT, ER, and TP, respectively, a study's results are comparable with most other 3D XR Fitts' studies. Besides these step sizes are suitable for each metric, e.g., a step size of 2 bits/s for TP can be greatly less informative and comparable than 0.5 bits/s, based on its nature and observed reported values in investigated studies.

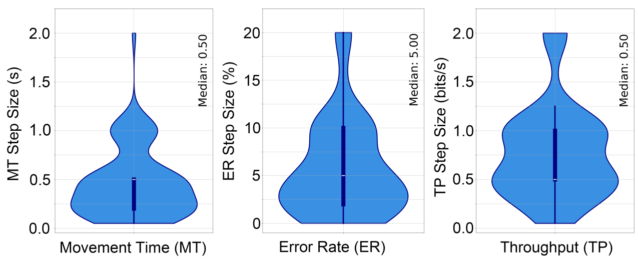

Figure 4: Density of different step sizes used in diagrams for reporting Movement Time (MT), Error Rate (ER), and Throughput (TP)

5.4.2 Trends Based on ID. We analyzed the average range for MT, ER, and TP based on the IDs. Among all studies, 80 (65.6%), 60 (49.2%), and 56 (45.9%) reported MT, TP, and ER along with IDs, respectively. Figure 5 represents the MT, ER, and TP ranges against used IDs. We plotted these measures to investigate how well the captured values in the studies match. For each measure, the lower and upper bounds of the reported range are mapped with the lower and higher bounds of the used IDs. If either the measure (MT, TP, ER) or ID was reported as a single value or average in a study, it is represented as a point on the plot. When one measure is a range and the other is a single value, the plot uses the midpoint of the range along with the single value to illustrate it with a dot.

As can be seen in Figure 5 (a), the area between MTs of 0.5s and 2s, and IDs of 2 and 4.5 bits, is notably denser. This density shows how reported MTs cluster between IDs of 2 to 4.5 bits. As for the ER (Figure 5 (b)), reported results cluster between 0% to 20% across mentioned IDs. In addition, although the common area of reported TPs in Figure 5 (c) is not as dense as the two others, it shows that most of the TPs reported in 3D XR Fitts' user studies are distributed between 1.5 and 5 bits/s. The IDs themselves, which are presented in Figure 5, show how the IDs are distributed between the ISO recommendation 2 and 8 bits. IDs between 2 and 4.5 are much more commonly used and never exceed 8 bits. These distributions indicate that our current knowledge is largely centered on relatively low-difficulty selection tasks, leaving more complex interactions at higher IDs underexplored.

5.4.3 Fitts' Regression. Fitts' regression is widely used in HCI studies, which is also a repeatedly used approach for showing the performance of users with linear regression of MT and ID. Our

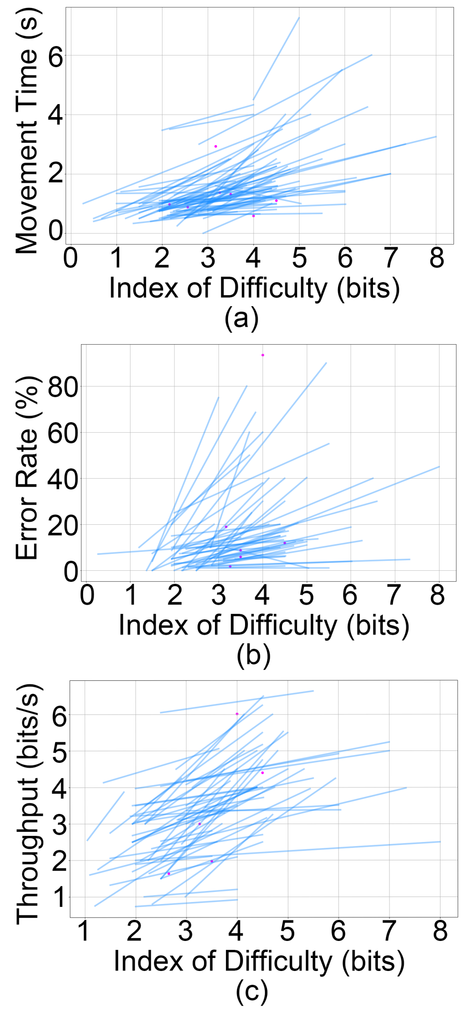

Figure 5: Plots representing the average ranges of movement time (a), error rate (b), and throughput (c) based on the reported IDs.

meta-analysis also investigated the distribution of reported slopes in studies, to represent the trend. Figure 6, shows the density and median of reported slopes (b coefficient in Fitts' law MT equation), representing the trend of captured values in 3D XR studies. A main inconsistency regarding Fitts' regression is the metric used for MT. 39.6% of studies used milliseconds and 60.4% seconds, resulting in variability and confusion in reports, especially if values are reported in a table instead of a plot. For instance, at a glance, a naive reader might interpret the results as vastly different (e.g., thinking Study A's MT is 1000 times higher than Study B's), when in fact they represent the same duration. This lack of clarity undermines the reliability of comparisons and can mislead further analysis or application of the findings. More concerningly, MT must be measured in seconds to ensure TP is correctly expressed in bits/s. If MT is mistakenly reported in milliseconds instead of seconds, the resulting TP value could be incorrect. Moreover, inconsistencies can

Table 4: Number and precision information for reported Movement Time (MT), Throughput (TP), and Error Rate (ER)

	MT (s)	TP (bits/s)	ER (%)
Number of Studies Measured	114 (93.4%)	84 (68.8%)	85 (69.7%)
Reported Exact Value	46.9%	54.2%	51.2%
Average Step Size (sd)	0.45 (0.36)	0.78 (0.48)	5.85 (5.17)

affect the linear regression analysis results, leading to larger a and b values. Our meta-analysis shows that 53.3% of studies used the Fitts' regression to report the relation between MT and ID in their study. This highlights the popularity of such regression as an evaluation tool in 3D XR. Nevertheless, among studies that reported regression, 26.1% presented the diagram without information related to the regression coefficient (especially slope), making cross-study of this important measure difficult.

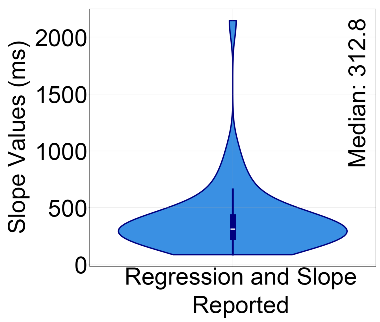

Figure 6: Density and median of reported slopes in studies that calculated Fitts' regressions.

5.5 Task Design

5.5.1 Target Placement. Fitts' law originally investigated 1D linear serial selections [52]. In XR, new target patterns are being explored, especially in terms of the third dimension, i.e., visual depth. Our analysis showed that the distribution of targets in Fitts' task is among the sources of variability in methodologies in 3D XR studies. Mainly there are 5 different designs used in experiment designs as listed in Table 5.

Among the approaches listed in Table 5, the circular distribution of targets is widely used in experiment designs, which also aligns with the extensive use of ISO 9241-411 multi-directional selection in XR studies. Our meta-analysis also revealed variability in designs and methodologies within each target distribution category. For example, the circular pattern, recommended by Soukoreff and MacKenzie [132] includes various implementations, e.g., Fernandes et al. [51] used a double-ISO design for their experimental task, resulting in a spectrum of task designs.

Another factor related to the target placement and task design analyzed during our analysis is the transparency and clarity in reporting A, W, and target Depth (D) in studies. A and W are the main attributes in the original Fitts' task design [52] that directly impact the index of difficulty in all extensions of Fitts' law. Of the studies reviewed, 84% reported A and W but in various metrics such as Euclidean distances (71.8%), angular distances (25.3%), and pixels (8.7%). In addition to the original A and W variables, D (depth) is also an important characteristic of 3D XR studies, indicating how far the target is placed from the screen surface or the user. Unfortunately, in 45.1% of the investigated studies, D is not explicitly indicated as part of the task design.

5.5.2 Feedback Type. Three main feedback types are used in 3D XR studies, visual (e.g., highlight or target change of color), auditory (e.g., beep or error sound), haptic feedback, which can be either active haptic (e.g., electrovibration [161] or passive haptics [101]). Many examples show that such feedback can affect the user's performance in selection tasks [19, 24, 101, 161]. Thus, many studies used one or a combination of these feedback methods in their experimental task. Our meta-analysis showed that visual feedback is the most commonly used feedback, with 121 (99.2%) studies reporting its use. One study, by Lin et al. [87], did not directly mention incorporated feedback as a part of the methodology, but based on the provided figure of the experiment it likely also used visual feedback, too. The widespread use of visual feedback is understandable given its ease of implementation and alignment with the Fitts' selection task, which inherently relies on clear indications of the next target. Further, it also exploits the fact that humans are generally visual-dominant. After visual feedback, auditory feedback with 31 instances (25.4%) and haptic feedback with 22 (18.0%) are in second and third place, respectively. Each feedback type was either a constant part of the experiment design or a factor under investigation in the study. In contrast, of studies that investigated feedback type as an independent variable, haptic feedback is the most studied feedback with 15 instances (12.3%) (e.g., [21, 101, 161]), followed by visual (e.g., [53, 90]) and auditory feedback (e.g., [19, 24]) with 7 (5.7%) and 6 (4.91%) studies respectively. Throughout the meta-analysis, we investigated the effect of feedback within Fitts' study results in 3D XR in Table 6. Figure 7 - Figure 9 show the density of reported values based on each feedback category, in different levels of usage (overall, feedback as the condition under study, and feedback constantly presents during the experiment).

To further illustrate how studies are clustered based on the feedback type incorporated, plotting the reported MT, TP, or ER based on used IDs is not informative, since the number of studies for each feedback category varies greatly. To address this, Figure 10 - Figure 12 illustrate the clusters of reported values based on IDs in different studies and visualize a correlation between the average results and the number of studies for each feedback type. In these figures, the average used ID and reported MT, TP, or ER is calculated (for studies that reported both), and illustrated as dots. The area of the bubble indicates the number of studies that used specific types of feedback and center crosses illustrate the overall

Table 5: List of different target distributions in Fitts' task designs.

#	Target Distribution Pattern	% of Studies	Examples
1	Circular	74.4 %	[10, 19, 20, 54]
2	Grid	8.2 %	[17, 22, 141]
3	1D Linear	7.4 %	[25, 146]
4	Distributed in 3D Space	7.4 %	[38, 163]
5	Bi-Directional (Lateral and/or In-Depth)	2.6 %	[15]

Table 6: Feedback types used in 3D XR Fitts' studies

#	Feedback Type	Identifier	% of Studies
1	Visual only	V	61.6%
2	Visual and Auditory	VA	20.4%
3	Visual and Haptic	VH	12.3%
4	Visual, Auditory, and Haptic	VAH	5.7%

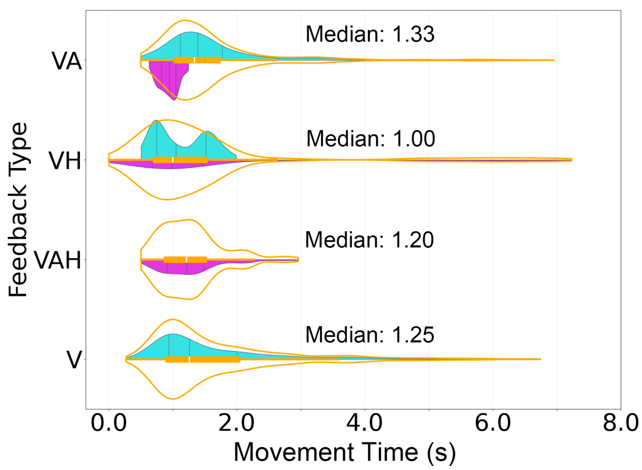

Figure 7: Density of reported movement times based on the used feedback in studies, for three levels of usage, feedback as a condition under study (pink), feedback present in all conditions of study (blue), and overall (yellow). The median is also listed.

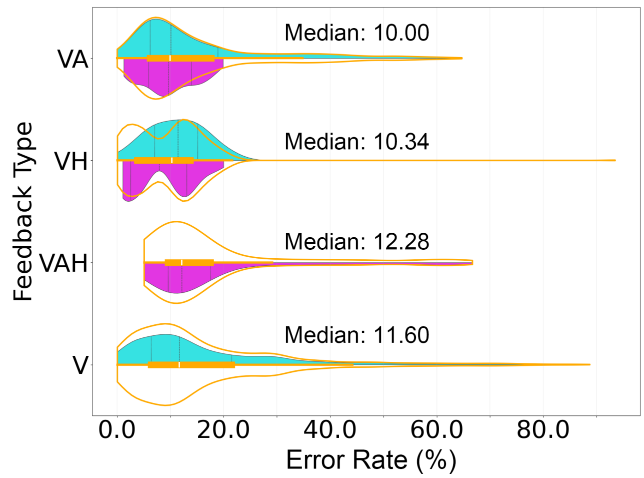

Figure 8: Density of reported error rates based on the used feedback in studies, for three levels of usage, feedback as a condition under study (pink), feedback present in all conditions of study (blue), and overall (yellow). The median is also listed.

average for each combination (bubbles are scaled by 8 for better readability).

As can be seen in Figure 10 - Figure 12, used IDs are mostly clustered between 2 to 4 bits for MT and 3 to 4 bits for ER and TP in different feedback categories. The sizes of the bubbles illustrate the portion of studies for each category, where visual only is the most reported one. Interestingly, it seems that the combination of visual and haptic feedback might (across all involved studies) exhibit overall higher throughput than visual alone, see Figure 12.

5.6 Participants

Since participants are the main source of data in a user study, we also analyzed participant attributes. Participant characteristics are fundamental to understanding the outcomes of a user study in HCI. This analysis may also inform future research to manage the corresponding participant selection.

5.6.1 Number and Age of Participants. Among the analyzed 3D XR studies, all reported the number of participants, and 117 (95.9%) reported the age of the participants. The mean age of participants was 25.81 (SD = 4.26). The number of participants varied between studies, but on average, experiments included 18 participants (SD = 8.52). Figure 13 illustrates the distribution of participant ages, showing both the density of ages with the violin plot and the summary statistics, including the median and quartiles, with the embedded box plot.

5.6.2 Gender of Participants. Of all the studies investigated, 110 (90.2%) reported the gender of participants, highlighting the importance of considering gender in participant selection. In total, 62.6% were male and 36.8% were female. In other words, slightly more than one-third of the participants were female, and a bit less

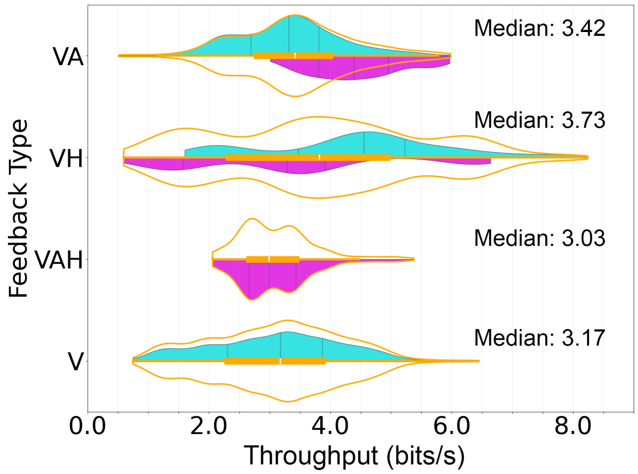

Figure 9: Density of reported throughputs based on the used feedback in studies, for three levels of usage, feedback as a condition under study (pink), feedback present in all conditions of study (blue), and overall (yellow). The median is also listed.

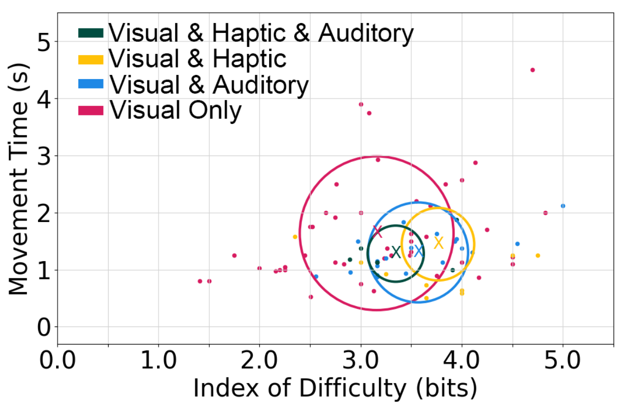

Figure 10: Average reported movement times based on the index of difficulty for different feedback combinations

than two-thirds of the participants were male. The gap between the percentage of male and female participants is due to the small number of studies that reported answers like non-binary, other, prefer not to say (e.g., [14, 51, 54, 55, 73, 122]). In total, 93.6% of studies solely differentiate between male and female, not reporting other genders.

6 Discussion

Fitts' law is widely used to investigate user interaction in 3D XR research. It has been widely applied in a diverse range of studies and applications within the XR field. Despite its significant impact in 2D HCI research, inconsistencies and variability in the application of Fitts' law within XR remain problematic [81, 145]. Our systematic review of Fitts' law in 3D XR reveals a research landscape that is both broad and fragmented. While each variation of

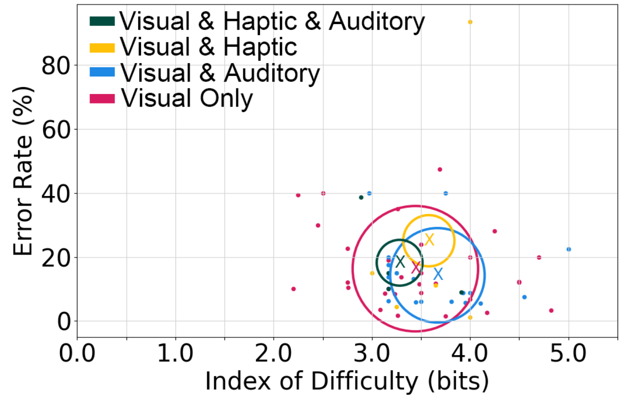

Figure 11: Average reported error rates based on the index of difficulty for different feedback combinations.

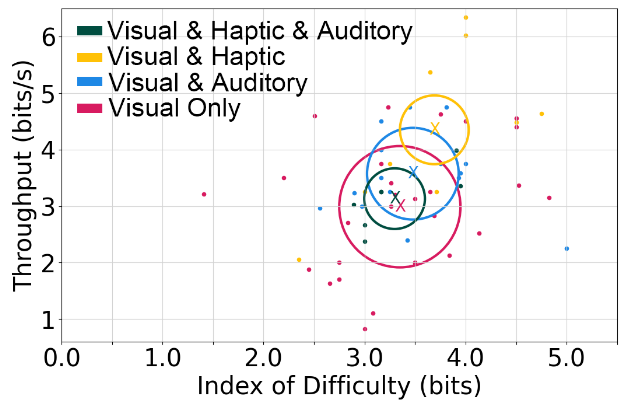

Figure 12: Average reported throughput based on the index of difficulty for different feedback combinations.

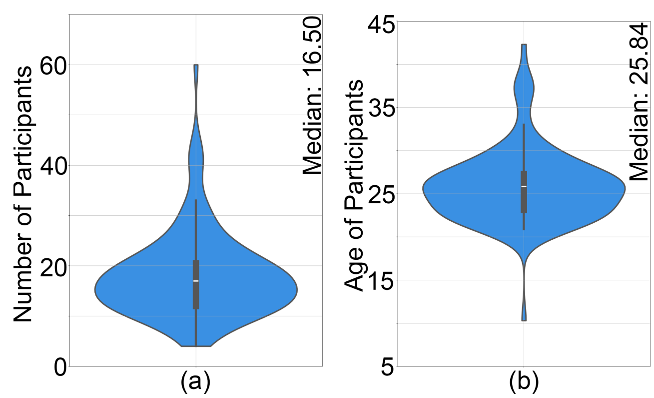

Figure 13: Distribution of the number of participants (a) and age of participants (b) in 3D XR Fitts' studies.

Fitts' law is designed to formulate movement time in a specific task environment, such as angular movements or targeting movements that change only the depth, these variations make it difficult for the research community to establish a standard approach. As a result, experiments using different versions of Fitts' law are often tailored to specific characteristics of XR interactions, such as target depth or the arrangement of objects in 3D space, whether planar or curved. There is no single formulation of Fitts' law satisfying all the different characteristics of XR interactions. Yet, as our focus is on the practices currently used in the field [58], whether a singular Fitts' law formulation should be established with an analysis of mathematical formulations is beyond the scope of this paper. Still, the focus of our systematic review is to address the lack of standardization, identify the common ground and sources of variability, and provide a framework to enable future research to move towards more consistent reports.

6.1 Variability and Inconsistency in 3D XR Fitts' Studies

Lack of Standard Approaches for Fitts' Law Studies in XR: Through our systematic review, we demonstrated the substantial influence and diverse applicability of Fitts' law in XR research. However, the lack of standard approaches for Fitts' law considerably hinders the reliability and comparability of results, blurring the global sense of outcomes in the research community. The very purpose of the standardized methodologies introduced originally in ISO 9241-9 [66] and later ISO 9241-411 [67] is to present a single consistent method for pointing device evaluation. These standards were adopted by the 2D HCI community in recognition of the importance of having a comparable methodology to support the replication of experiments, a key tenet of scientific research across all disciplines, from which HCI is not exempt [96]. This consistency has been instrumental in facilitating the growth of the research field on 2D selection interfaces, and we argue it is similarly important in the 3D XR domain. For instance, using different Fitts' law extensions provides different levels of model accuracy (R^2) in 3D target acquisition task [70].

Various Experimental Setups and Tasks: Different experiment setups and task designs can influence users' performance, e.g., if targets are highlighted [33, 138], whether selection errors are permissible (e.g., [15, 19, 144]) or not (e.g., [104, 108, 156]), the direction of selections [108], the position of the targets (as selection in depth is less accurate and slower [138]), or considerations regarding the Heisenberg spatial effect [29]. For instance, while [153] and [164] used the same interaction technique and feedback with an HMD and overlapping ID ranges ([2.82–3.70] and [1.37–3.64], respectively), their TP results did not overlap ([2.49–3.11] and [4.13–5.06]). This discrepancy may be attributed to differences in target depths or experimental methodologies but is inconclusive due to a lack of standardized methodologies and reporting. This is one of the inherent problems of the use of Fitts' law in 2D research [58]. Although ISO 9241 [66, 67, 94] was a great step toward the standardization of methods, with the rapidly growing research in 3D XR, the clarity of description and solid justifications should not be neglected. Our systematic review shows that with more complex characteristics of XR interactions, like proximal (e.g., [15, 17]) and distal interactions (e.g., [18, 146]), correct and more accurate application of Fitts' law plays a crucial role in maintaining a high quality of research. ***It is important to note that simply mentioning Fitts' law or using the movement time formulation to validate results is not necessarily sufficient, particularly if a study aims to understand and compare user performance.***

Fitts' Law Variations: The main sources of variability in 3D XR Fitts' studies are the use of diverse variations of the original model, the different types of displays and interaction methods explored, the multiple task designs with different approaches towards target placement, and the use of heterogeneous metrics. These inconsistencies, from task design to result reporting, exacerbate the challenges posed by the lack of a standard, particularly as Fitts' law applications in XR continue to grow rapidly. Addressing these issues is necessary for creating a climate of effective future research, intending to limit the blind growth of variability, and to provide better options for the evaluation of XR real-world contexts.

6.2 Trends in 3D XR Fitts' Studies

Performance Metrics and Experimental Trends: Our meta-analysis shows the trend of different measures related to Fitts' experiments, particularly movement time which is the most studied measure, error rate, and throughput. Selection tasks in XR headsets take 1.25 seconds, with an error rate of 11.95% and throughput of 3.25 bits/s, though these values vary depending on feedback, chosen IDs, and participant demographics. Moreover, the average age of the participants is 28, and 36.8% of them female, 62.6% of them male, and 0.6% of them reported as binary, other, or prefer not to say. The majority of the papers used Shannon formulation Equation 6 or Effective Fitts' law equation Equation 8 (78.5%), which is also used in the ISO 9241 document. The majority of the papers (62.3%) used Fitts' law for analysis of interaction and system evaluation. Our results also revealed that most user studies were conducted with an HMD (83.5%). Overall, investigated papers used a circular target distribution pattern (78%) for user performance analysis. The majority of the papers reported that they only used visual feedback (61%). Moreover, in Appendix Table 7, we summarized Fitts' law equation variations for 3D XR for different interaction spaces, displays, interaction techniques, and target types that are commonly used.

Implications in 3D XR Studies: Our detailed meta-analysis revealed the wide distribution of reported values in 3D XR research. Reported clusters highlight the overall trends and the current state of results, such as feedback type, and the range of IDs explored across studies. The observed clusters not only reflect how different experimental designs and task complexities impact the measured outcomes but also highlight the variability in methodologies and their implications on performance metrics. For instance, MacKenzie [94] recommended using IDs between 2 to 8 bits for HCI user experiments. Our meta-analysis shows that used IDs in 3D XR studies not only lie within this range but also do not exceed a difficulty of 8 bits, with the majority of the studies investigating the user performance within the range of 2 to 4.5 bits. This is expected, as, for instance, the accuracy of a 2D tabletop mouse and a human hand hovering in mid-air have different ergonomic limitations associated.

A task with a difficulty of 8 bits is challenging for virtual hand interaction since it requires a high level of precision and control. Besides, the synthesized findings of reported values in 3D XR research can be used as a tool for further assessment of the results of future user studies. As an example, if a study reports values around 3 to 5 s as the movement time for a task with the index of difficulty equal to 4 bits, this very likely indicates either a sub-optimal interaction technique or that the study is suffering from some other issue that requires the work to be re-investigated.

6.3 Recommendations and Insights

Although Fitts' law serves as a common conceptual foundation, the field currently lacks consistent methodologies, as discussed above, standardized reporting practices, and clear guidelines for adapting the law to the complexities of XR. These issues manifest themselves in multiple ways. For instance, while many studies rely on effective-measures formulations or variations such as Kopper et al.'s angular equation (Equation 11), the choice of the formula is rarely justified, and sometimes even unclear. This situation hinders meaningful comparisons across studies. Moreover, measurements crucial for reproducibility, such as ID, MT, ER, and TP,—are often reported with insufficient precision or inconsistent units. A and W are measured in everything from angular degrees to pixels, while MT might be reported in seconds or milliseconds without explicit mention. Variability in reporting feedback modalities (visual, auditory, haptic), task design (circular, grid-based, linear, or fully 3D distributions), and participant demographics further complicates efforts to synthesize results. This harms the HCI community as it undermines the consistency and comparability of the studies.

Additionally, although over 90% of studies report the gender distribution of participants, the data remain skewed, with female and non-binary participants underrepresented. The lack of demographic diversity—and incomplete reporting thereof—could limit our understanding of how different populations experience and perform in 3D XR tasks.

Even though consistency is instrumental in facilitating the growth of research, there has been no silver bullet in 3D XR Fitts' law studies. Taken together, these observations signal a pressing need to move beyond mere "documentation" of what has been done. Instead, the field should coalesce around clearer, well-rationalized methodologies, provide more transparency in reporting, and adopt standardized frameworks that facilitate cross-study comparisons. The goal is to empower researchers and practitioners to build on each other's work more effectively and to ensure that conclusions are robust, equitable, and broadly applicable.

Our framework for conducting Fitts' law studies, especially in 3D XR, is provided in Appendix Table 8, as are its practical implications that can be used by researchers, developers, and practitioners. We also hope that reviewers can use this document to evaluate future 3D XR papers that use Fitts' law as a methodology. Below, we summarized this table with accessible approaches. Rather than simply stating that standardized methods are needed, we detail what those methods might look like, who could define them, and how researchers can adopt them.

Adopt and Reference Standardized Frameworks: To improve the quality and consistency of Fitts' law research in 3D XR, adopting standardized frameworks is crucial. Established guidelines, such as ISO 9241-411 [67], should be referenced to ensure adherence to commonly accepted methodologies. Researchers should explicitly state which version of Fitts' law they employ—such as the effective-width measures—while including the exact equation, units used, and rationale for the chosen variation. By clearly referencing these standards, studies can better align with benchmarks, enabling improved cross-study comparability and fostering a cohesive research framework.

Use Consistent Units and Scales in Reporting: Consistency in units and scales is another essential practice. Defining a default unit system, such as centimeters or degrees for distances and seconds for time, would greatly enhance clarity. Recommended granularity, such as increments of 0.5 bits/s for throughput (TP) or 5% for error rate (ER), can facilitate direct comparisons. Before collecting data, researchers should decide on and document all units, ensuring MT is consistently reported in seconds (not milliseconds) and target dimensions are converted into commonly understood metrics like centimeters or degrees (e.g., not pixels). The uniform use of units and scales minimizes confusion and allows researchers to integrate results across studies more effectively.

Standardize Task Designs and ID Selections: Standardization of task designs and ID selections can further streamline research. It is essential to commit to commonly used ID ranges, such as 2 to 4.5 bits, and to document the target amplitude (A) and width (W) thoroughly. This documentation should include details on how these values are derived, whether through Euclidean or angular distances, with clear justification. A simple reference table summarizing IDs, A, W, and depth (D) for each condition should be provided in publications. Specifying whether A and W are effective or nominal and rationalizing the chosen metrics, such as degrees for ray casting or centimeters for virtual hand interactions, ensures that task parameters are more uniform and studies are easier to replicate and compare.

Guidelines for Feedback Modalities and Reporting: Researchers should provide comprehensive information on the type, intensity, duration, and rationale for feedback used in experiments. For instance, if haptic feedback is introduced, the device's technical parameters and calibration methods should be specified. Similarly, auditory feedback should include volume and frequency ranges. Transparent documentation of feedback conditions not only enables accurate replication of the study but also helps identify how different modalities influence user performance.

Justify Methodological Choices and Clearly Report Regression Models: When employing regression analyses—such as linear fits between MT and ID—it is critical to report the slope, intercept, goodness of fit, and units used with clarity. A supplementary table summarizing regression parameters for each condition can improve transparency. Additionally, researchers should explicitly state any unit conversions. Clear reporting of regression models ensures future researchers can accurately interpret and integrate findings into meta-analyses without ambiguity.

Improve Transparency in Participant Demographics: Transparency in participant demographics is another key consideration. Studies should move beyond a binary male/female distinction and clearly report non-binary or unreported genders. Researchers should include a rationale for participant demographic selections and discuss their implications for the generalizability of findings. Inclusive recruitment strategies can be enhanced by providing a demographics table with categories such as male, female, non-binary, other, and prefer not to say. Encouraging self-identification rather than imposing predefined categories ensures more representative samples.

Pre-Study Protocol Registration and Data Sharing Sharing data can significantly enhance research transparency [118]. Researchers should consider open-source platforms to share their methodologies, including chosen ID levels, feedback details, and demographic plans. By making anonymized raw datasets available, future researchers can replicate or reanalyze results, thereby strengthening the collective reliability of findings in the field. These practices ensure that Fitts' law research in 3D XR continues to evolve in a rigorous and reproducible manner.

We also recommend that researchers explain and give details of their study if they cannot follow these recommendations. Since Fitts' law is used to compare different input and output devices and interaction methods, the existing framework might not address all the challenges. For instance, Batmaz et al.[23] changed the layout of the ISO 9241:411 multidirectional selection task. In such cases, we recommend providing a clear justification for the modifications made, including detailed descriptions of the study design, task adjustments, and their rationale. This ensures transparency and allows other researchers to replicate or build upon the work, contributing to the broader understanding of Fitts' law applications in novel contexts.

By transitioning from ad-hoc, case-by-case methodologies towards more standardized and transparent frameworks, future Fitts' law research in 3D XR can achieve greater consistency, comparability, and relevance. These recommendations are not mere suggestions: they outline a path forward, guiding researchers to make deliberate, well-documented choices at each stage of their study. Such practices will accelerate progress in understanding how users interact with XR systems, enabling researchers, developers, and practitioners to design more effective and inclusive environments.

6.4 Open Problems and Considerations

Through our identification of the trends and gaps in 3D XR Fitts' studies, we present an overview of open problems that can be addressed and further investigated by the research community. For instance, future research should examine user performance across a wider range of IDs to better understand how task difficulty affects performance, particularly for high-difficulty tasks. These tasks, often explored in desktop environments, remain underexplored in XR contexts. Addressing such gaps would contribute to a more comprehensive and robust application.

Moreover, more studies about the effect of sources of variability in Fitts' tasks might help the research community as a whole to move towards more standardized evaluations, ultimately bringing all the currently dispersed results for the main interaction methods (e.g., ray casting and virtual hand) together to directly affect the real-world application of XR. As an example, future research efforts can explore the effect of feedback type on user performance with a focus on Fitts' task or, in particular, different combinations of feedback types. Our synthesis shows that there is less focus on incorporating multiple feedback types together in Fitts' studies, such as visual and haptic (12.3% of studies reported such feedback), or visual and haptic with auditory feedback (5.7%), all opportunities for further exploration.

While standards such as ISO 9241:411 provide valuable guidelines for Fitts' law studies, their suitability and efficiency for 3D XR contexts remain largely untested. These standards were originally developed for 2D interaction environments, and there is no empirical evidence demonstrating whether they adequately account for the unique complexities of 3D XR interactions, such as depth perception, spatial orientation, or multimodal feedback. For instance, our review showed that there is more focus on the circular arrangement of targets in Fitts' study (e.g., [19, 20]). However, our results indicate that target arrangements are mostly placed at a constant depth distance, which limits the performance assessment of the analyzed methodology or interaction technique. We also recommend evaluating the validity of the methodologies with targets at different depth distances by using the appropriate Fitts' law formulation. In other words, it is essential to conduct empirical analyses to determine whether these models need to be adjusted or extended to address the challenges posed by 3D XR environments. Such investigations would provide clarity on the applicability of existing standards and inform the development of optimized frameworks for 3D XR studies. It is worth mentioning that Fitts' law is proposed considering pointing tasks with rapid aimed movements, creating a balance in speed-accuracy trade-off [52, 94]. However, it should not be considered as the *only* method for evaluating XR systems or user performance. Especially, owing to growing interactions in 3D XR, other alternative methods, e.g., "targets inside a volume" for dense environments [13], can be used to investigate user performance and interaction techniques.

This systematic review mainly focuses on Fitts' law studies in 3D XR, in particular, for target pointing and selection tasks. However, Fitts' law extends beyond this scope and applies to other tasks, including those modeled by extensions like the steering law [1].

7 Conclusions

In conclusion, our systematic review and meta-analysis indicate the variable application of and diverse approaches used for Fitts' law in 3D XR research, and emphasize the importance of precise reporting and thoughtful task design in Fitts' law studies within 3D XR environments. The findings reveal that while Fitts' law provides a robust framework for evaluating and designing interaction techniques, variability in the variation of the Fitts' law model used in the study, many aspects of the experimental task, and the metric precision of the results reported can influence study outcomes and their comparability. Our analysis highlights the predominant focus on specific demographics in current research and points to opportunities for expanding studies to include more diverse demographics. By adhering to the recommendations provided and clearly reporting which Fitts' law-related methodology was used, researchers can enhance the reliability and applicability of their findings, ultimately contributing to a clearer understanding of the many subtle aspects of interaction performance in the 3D XR context. Overall, as the field continues to evolve, our work will help researchers shape the direction of future research, ensuring that it builds on a solid foundation of past knowledge.

Acknowledgments

This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, award number: RGPIN-2023-05642.

References

[1] Johnny Accot and Shumin Zhai. 1997. Beyond Fitts' law: models for trajectory-based HCI tasks. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*. 295–302.

[2] Johnny Accot and Shumin Zhai. 2003. Refining Fitts' law models for bivariate pointing. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 193–200.

[3] ACM. 2004. ACM Digital Library. https://dl.acm.org. Last accessed: 2024-08-01.

[4] Riaz Ahmed Agha, Alexander J Fowler, Christopher Limb, Katharine Whitehurst, Robert Coe, Harkiran Sagoo, Daniyal J Jafree, Charmilie Chandrakumar, and Buket Gundogan. 2016. Impact of the mandatory implementation of reporting guidelines on reporting quality in a surgical journal: a before and after study. *International journal of surgery* 30 (2016), 169–172.

[5] Hirotugu Akaike. 1974. A new look at the statistical model identification. *IEEE transactions on automatic control* 19, 6 (1974), 716–723.

[6] Ahmed Alnagrat, Rizalafande Che Ismail, Syed Zulkarnain Syed Idrus, and Rawad Mansour Abdulhafith Alfaqi. 2022. A review of extended reality (XR) technologies in the future of human education: Current trend and future opportunity. *Journal of Human Centered Technology* 1, 2 (2022), 81–96.

[7] Ahmed Jamah Ahmed Alnagrat, Rizalafande Che Ismail, and Syed Zulkarnain Syed Idrus. 2021. Extended reality (XR) in virtual laboratories: A review of challenges and future training directions. In *Journal of Physics: Conference Series*, Vol. 1874. IOP Publishing, 012031.

[8] Farshid Amirabdollahian, Germano T Gomes, and Garth R Johnson. 2005. The peg-in-hole: a VR-based haptic assessment for quantifying upper limb performance and skills. In *9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005*. IEEE, 422–425.

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

[10] Oscar Ariza, Gerd Bruder, Nicholas Katzakis, and Frank Steinicke. 2018. Analysis of proximity-based multimodal feedback for 3d selection in immersive virtual environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 327–334.

[11] Frédéric Banville, Pierre Nolin, Thaïna Rosinvil, Eulalie Verhulst, and Philippe Allain. 2019. Assessment and rehabilitation after traumatic brain injury using virtual reality: A systematic review and discussion concerning human-computer interactions. *Virtual reality for psychological and neurocognitive interventions* (2019), 327–360.

[12] Mayra Donaji Barrera Machuca and Wolfgang Stuerzlinger. 2019. The effect of stereo display deficiencies on virtual hand pointing. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–14.

[13] Mohammad Raihanul Bashar and Anil Ufuk Batmaz. 2024. Virtual Task Environments Factors Explored in 3D Selection Studies. In *Proceedings of the 50th Graphics Interface Conference*. 1–16.

[14] Anil Ufuk Batmaz, Mayra Donaji Barrera Machuca, Junwei Sun, and Wolfgang Stuerzlinger. 2022. The effect of the vergence-accommodation conflict on virtual hand pointing in immersive displays. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15.

[15] Anil Ufuk Batmaz, Mayra Donaji Barrera Machuca, Duc Minh Pham, and Wolfgang Stuerzlinger. 2019. Do head-mounted display stereo deficiencies affect 3D pointing tasks in AR and VR?. In *2019 IEEE conference on virtual reality and 3D user interfaces (VR)*. IEEE, 585–592.

[16] Anil Ufuk Batmaz, Moaaz Hudhud Mughribi, Mine Sarac, Mayra Barrera Machuca, and Wolfgang Stuerzlinger. 2023. Measuring the effect of stereo deficiencies on peripersonal space pointing. In *2023 IEEE conference virtual reality and 3D user interfaces (VR)*. IEEE, 1–11.

[17] Anil Ufuk Batmaz, Aunnoy K Mutasim, Morteza Malekmakan, Elham Sadr, and Wolfgang Stuerzlinger. 2020. 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)*. IEEE, 184–193.

[18] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2019. Effects of 3D rotational jitter and selection methods on 3D pointing tasks. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 1687–1692.

[19] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2021. 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)*. IEEE, 85–94.

[20] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2022. Effective throughput analysis of different task execution strategies for mid-air fitts' tasks in virtual reality. *IEEE Transactions on Visualization and Computer Graphics* 28, 11 (2022), 3939–3947.

[21] Anil Ufuk Batmaz, Xintian Sun, Dogu Taskiran, and Wolfgang Stuerzlinger. 2019. Hitting the wall: Mid-air interaction for eye-hand coordination. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology*. 1–5.

[22] Anil Ufuk Batmaz, Xintian Sun, Dogu Taskiran, and Wolfgang Stuerzlinger. 2020. Eye-hand coordination training for sports with mid-air vr. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*. 1–10.

[23] Anil Ufuk Batmaz, Rumeysa Turkmen, Mine Sarac, Mayra Donaji Barrera Machuca, and Wolfgang Stuerzlinger. 2023. Re-investigating the effect of the vergence-accommodation conflict on 3d pointing. In *Proceedings of the 29th ACM symposium on virtual reality software and technology*. 1–10.

[24] Anil Ufuk Batmaz, Kangyou Yu, Hai-Ning Liang, and Wolfgang Stuerzlinger. 2022. Improving effective throughput performance using auditory feedback in virtual reality. In *Proceedings of the 2022 ACM Symposium on Spatial User Interaction*. 1–11.

[25] Joshua D Bell and Kristen L Macuga. 2019. Goal-directed aiming under restricted viewing conditions with confirmatory sensory feedback. *Human movement science* 67 (2019), 102515.

[26] Joanna Bergström, Aske Mottelson, Andreea Muresan, and Kasper Hornbæk. 2019. Tool extension in human-computer interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–11.

[27] Olav Bertelsen. 1994. Fitts' law as a design artefact: A paradigm case of theory in software design. In *International Conference on Human-Computer Interaction*. Springer, 11–18.

[28] Jae Hwan Bong, Sunwoong Choi, Jin Hong, and Shinsuk Park. 2022. Force feedback haptic interface for bilateral teleoperation of robot manipulation. *Microsystem Technologies* 28, 10 (2022), 2381–2392.

[29] Doug Bowman, Chadwick Wingrave, Joshua Campbell, and Vinh Ly. 2001. Using pinch gloves (tm) for both natural and abstract interaction techniques in virtual environments. (2001).

[30] Doug A Bowman and Larry F Hodges. 1997. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proceedings of the 1997 symposium on Interactive 3D graphics*. 35–ff.

[31] Doug A Bowman, Donald B Johnson, and Larry F Hodges. 1999. Testbed evaluation of virtual environment interaction techniques. In *Proceedings of the ACM symposium on Virtual reality software and technology*. 26–33.

[32] John Brooke. 2013. SUS: a retrospective. *Journal of usability studies* 8, 2 (2013).

[33] Gerd Bruder, Frank Steinicke, and Wolfgang Sturzlinger. 2013. To touch or not to touch? comparing 2D touch and 3D mid-air interaction on stereoscopic tabletop surfaces. In *Proceedings of the 1st Symposium on Spatial User Interaction* (Los Angeles, California, USA) *(SUI '13)*. Association for Computing Machinery, New York, NY, USA, 9–16. https://doi.org/10.1145/2491367.2491369

[34] Kenneth P Burnham, David R Anderson, Kenneth P Burnham, and David R Anderson. 1998. *Practical use of the information-theoretic approach*. Springer.

[35] Stuart K Card, William K English, and Betty J Burr. 1978. Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. *Ergonomics* 21, 8 (1978), 601–613.

[36] Juan Sebastián Casallas, James H Oliver, Jonathan W Kelly, Frédéric Merienne, and Samir Garbaya. 2013. Towards a model for predicting intention in 3D moving-target selection tasks. In *Engineering Psychology and Cognitive Ergonomics. Understanding Human Cognition: 10th International Conference, EPCE 2013, Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013, Proceedings, Part I 10*. Springer, 13–22.

[37] Yeonjoo Cha and Rohae Myung. 2013. Extended Fitts' law for 3D pointing tasks using 3D target arrangements. *International Journal of Industrial Ergonomics* 43, 4 (2013), 350–355.

[38] Yuan Chen, Keiko Katsuragawa, and Edward Lank. 2020. Understanding viewport-and world-based pointing with everyday smart devices in immersive augmented reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[39] Yi Fei Cheng, Tiffany Luong, Andreas Rene Fender, Paul Streli, and Christian Holz. 2022. ComforTable user interfaces: Surfaces reduce input error, time, and exertion for tabletop and mid-air user interfaces. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 150–159.

[40] Logan D Clark, Aakash B Bhagat, and Sara L Riggs. 2020. Extending Fitts' law in three-dimensional virtual environments with current low-cost virtual reality technology. *International Journal of Human-Computer Studies* 139 (2020), 102413.

[41] Edward Robert FW Crossman and PJ Goodeve. 1983. Feedback control of hand-movement and Fitts' Law. *The Quarterly Journal of Experimental Psychology* 35, 2 (1983), 251–278.

[42] Vernon R Curran, Xiaolin Xu, Mustafa Yalin Aydin, and Oscar Meruvia-Pastor. 2023. Use of extended reality in medical education: an integrative review. *Medical Science Educator* 33, 1 (2023), 275–286.

[43] Nguyen-Thong Dang. 2007. A survey and classification of 3D pointing techniques. In *2007 IEEE international conference on research, innovation and vision for the future*. IEEE, 71–80.

[44] Elena De Momi, Mahdi Tavakoli, Jeffrey Delmerico, Antonio Frisoli, Mark A Minor, Giovanni Rossini, and Paul Chippendale. 2022. Extended Reality in Robotics [From the Guest Editors]. *IEEE Robotics & Automation Magazine* 29, 1 (2022), 8–9.

[45] Sandrine De Ribaupierre, Ryan Armstrong, Dayna Noltie, Matt Kramers, and Roy Eagleson. 2015. VR and AR simulator for neurosurgical training. In *2015 IEEE Virtual Reality (VR)*. IEEE, 147–148.

[46] Ruth A Delaney. 2023. Extended reality technologies and their applications in shoulder replacement. In *Seminars in Arthroplasty: JSES*, Vol. 33. Elsevier, 839–845.

[47] Andreas Dünser, Martin Lochner, Ulrich Engelke, and David Rozado Fernández. 2015. Visual and manual control for human-robot teleoperation. *IEEE computer graphics and applications* 35, 3 (2015), 22–32.

[48] Elsevier. 2004. Elsevier ScienceDirect. https://www.sciencedirect.com. Last accessed: 2024-08-01.

[49] Orlando Erazo, José A Pino, Rosario Pino, Alfonso Asenjo, and Carmen Fernández. 2014. Magic mirror for neurorehabilitation of people with upper limb dysfunction using kinect. In *2014 47th Hawaii International Conference on System Sciences*. IEEE, 2607–2615.

[50] Augusto Esteves, Yonghwan Shin, and Ian Oakley. 2020. Comparing selection mechanisms for gaze input techniques in head-mounted displays. *International Journal of Human-Computer Studies* 139 (2020), 102414.

[51] Ajoy S Fernandes, T Scott Murdison, and Michael J Proulx. 2023. Leveling the playing field: A comparative reevaluation of unmodified eye tracking as an input and interaction modality for VR. *IEEE Transactions on Visualization and Computer Graphics* 29, 5 (2023), 2269–2279.

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

[53] Michael J Fu, Andrew D Hershberger, Kumiko Sano, and M Cenk Çavuşoğlu. 2011. Effect of visuo-haptic co-location on 3D Fitts' task performance. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 3460–3467.

[54] Jenny Gabel, Susanne Schmidt, Oscar Ariza, and Frank Steinicke. 2023. Redirecting Rays: Evaluation of Assistive Raycasting Techniques in Virtual Reality. In *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*. 1–11.

[55] Kristen Grinyer and Robert Teather. 2023. Low-Fi VR Controller: Improved Mobile Virtual Reality Interaction via Camera-Based Tracking. In *Proceedings of the 2023 ACM Symposium on Spatial User Interaction*. 1–12.

[56] Tovi Grossman and Ravin Balakrishnan. 2004. Pointing at trivariate targets in 3D environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 447–454.

[57] Jérôme Guzzi, Gabriele Abbate, Antonio Paolillo, and Alessandro Giusti. 2022. Interacting with a conveyor belt in virtual reality using pointing gestures. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 1194–1195.

[58] Chris Hand. 1997. A survey of 3D interaction techniques. In *Computer graphics forum*, Vol. 16. Wiley Online Library, 269–281.

[59] SG Hart. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload/Elsevier* (1988).

[60] Errol R Hoffmann. 1995. Effective target tolerance in an inverted Fitts task. *Ergonomics* 38, 4 (1995), 828–836.

[61] Dominic E Holmes, Darryl K Charles, Philip J Morrow, S McClean, and SM McDonough. 2016. Using fitt's law to model arm motion tracked in 3D by a leap motion controller for virtual reality upper arm stroke rehabilitation. In *2016 IEEE 29th international symposium on computer-based medical systems (CBMS)*. IEEE, 335–336.

[62] Matt C Howard and Maggie M Davis. 2022. A meta-analysis and systematic literature review of mixed reality rehabilitation programs: Investigating design characteristics of augmented reality and augmented virtuality. *Computers in Human Behavior* 130 (2022), 107197.

[63] Xuhui Hu, Aiguo Song, Zhikai Wei, and Hong Zeng. 2022. Stereopilot: A wearable target location system for blind and visually impaired using spatial audio rendering. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 30 (2022), 1621–1630.

[64] Jin Huang, John J Dudley, Stephen Uzor, Dong Wu, Per Ola Kristensson, and Feng Tian. 2022. Understanding user performance of acquiring targets with motion-in-depth in virtual reality. *International Journal of Human-Computer Studies* 163 (2022), 102817.

[65] IEEE. 2004. IEEE Xplore. https://ieeexplore.ieee.org. Last accessed: 2024-08-01.

[66] International Organization for Standardization 2000. *ISO 9241-9: Ergonomics of human-system interaction – Part 9: Requirements for non-keyboard input devices*. International Organization for Standardization. ISO Standard.

[67] International Organization for Standardization 2012. *ISO 9241-411:2012 Ergonomics of human-system interaction – Part 400: Evaluation methods for design of physical input devices*. International Organization for Standardization. ISO Standard.

[68] Nancy Jacho-Guanoluisa, Sergio Albiol-Pérez, Cesar A Collazos, Habib M Fardoun, and Sandra Cano. 2019. Cognitive rehabilitation using virtual reality for children with rare diseases. In *Proceedings of the 5th Workshop on ICTs for improving Patients Rehabilitation Research Techniques*. 183–186.

[69] Henri Jalo, Henri Pirkkalainen, Osku Torro, Elena Pessot, Andrea Zangiacomi, and Aleksei Tepljakov. 2022. Extended reality technologies in small and medium-sized European industrial companies: level of awareness, diffusion and enablers of adoption. *Virtual Reality* 26, 4 (2022), 1745–1761.

[70] Izabelle Janzen, Vasanth K Rajendran, and Kellogg S Booth. 2016. Modeling the impact of depth on pointing performance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 188–199.

[71] Rong Jiang and Zhenyu Gu. 2020. Current theoretical developments and applications of Fitts' Law: a literature review. In *Advances in Ergonomics in Design: Proceedings of the AHFE 2019 International Conference on Ergonomics in Design, July 24-28, 2019, Washington DC, USA 10*. Springer, 753–760.

[72] Kyungmo Jung, Baeksuk Chu, Shinsuk Park, and Daehie Hong. 2013. An implementation of a teleoperation system for robotic beam assembly in construction. *International Journal of Precision engineering and manufacturing* 14 (2013), 351–358.

[73] Kim Kargut, Carl Gutwin, and Andy Cockburn. 2024. Effects of Device Environment and Information Layout on Spatial Memory and Performance in VR Selection Tasks. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–17.

[74] Kibum Kim and Xiangshi Ren. 2014. Assisting visually impaired people to acquire targets on a large wall-mounted display. *Journal of Computer Science and Technology* 29, 5 (2014), 825–836.

[75] Woojoo Kim, Junyoung Jung, and Shuping Xiong. 2020. VRMouse: Mouse emulation with the VR controller for 2D selection in VR. In *Advances in Usability, User Experience, Wearable and Assistive Technology: Proceedings of the AHFE 2020 Virtual Conferences on Usability and User Experience, Human Factors and Assistive Technology, Human Factors and Wearable Technologies, and Virtual Environments and Game Design, July 16-20, 2020, USA*. Springer, 663–670.

[76] Luv Kohli, Mary C Whitton, and Frederick P Brooks. 2012. Redirected touching: The effect of warping space on task performance. In *2012 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 105–112.

[77] Yong-Ku Kong, Sang-Soo Park, Jin-Woo Shim, Kyeong-Hee Choi, Hyun-Ho Shim, Kiana Kia, and Jeong Ho Kim. 2023. A passive upper-limb exoskeleton reduced muscular loading during augmented reality interactions. *Applied ergonomics* 109 (2023), 103982.

[78] Regis Kopper, Doug A Bowman, Mara G Silva, and Ryan P McMahan. 2010. A human motor behavior model for distal pointing tasks. *International journal of human-computer studies* 68, 10 (2010), 603–615.

[79] Ka-Wai Kwok, Loi-Wah Sun, George P Mylonas, David RC James, Felipe Orihuela-Espina, and Guang-Zhong Yang. 2012. Collaborative gaze channelling for improved cooperation during robotic assisted surgery. *Annals of biomedical engineering* 40 (2012), 2156–2167.

[80] Fabien Lareyre, Claude Mialhe, Bahaa Nasr, Elise Poggi, Gilles Di Lorenzo, Khalid Rajhi, Arindam Chaudhuri, and Juliette Raffort. 2024. Extended and Augmented Reality in Vascular Surgery: opportunities and challenges. In *Seminars in Vascular Surgery*. Elsevier.

[81] Joseph J LaViola Jr, Ernst Kruijff, Ryan P McMahan, Doug Bowman, and Ivan P Poupyrev. 2017. *3D user interfaces: theory and practice*. Addison-Wesley Professional.

[82] Daniel A Lazewatsky, Cameron Bowie, William Curran, Jasper LaFortune, Benjamin Narin, Duy Nguyen, Amy Wyman, and William D Smart. 2014. Wearable computing to enable robot microinteractions. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 141–146.

[83] Victoria Leclercq, Charlotte Beaudart, Sara Ajamieh, Véronique Rabenda, Ezio Tirelli, and Olivier Bruyere. 2019. Meta-analyses indexed in PsycINFO had a better completeness of reporting when they mention PRISMA. *Journal of Clinical Epidemiology* 115 (2019), 46–54.

[84] Lik Hang Lee, Kit Yung Lam, Tong Li, Tristan Braud, Xiang Su, and Pan Hui. 2019. Quadmetric optimized thumb-to-finger interaction for force assisted one-handed text entry on mobile headsets. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–27.

[85] Nianlong Li, Teng Han, Feng Tian, Jin Huang, Minghui Sun, Pourang Irani, and Jason Alexander. 2020. Get a grip: Evaluating grip gestures for vr input using a lightweight pen. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[86] Yang Li, Sayan Sarcar, Kibum Kim, Huawei Tu, and Xiangshi Ren. 2022. Designing successive target selection in virtual reality via penetrating the intangible

interface with handheld controllers. *International Journal of Human-Computer Studies* 165 (2022), 102835.

[87] Chiuhsiang Joe Lin, Tien-Lung Sun, Hung-Jen Chen, and Ping-Yun Cheng. 2009. Evaluation of visually-controlled task performance in three dimension virtual reality environment. In *Virtual and Mixed Reality: Third International Conference, VMR 2009, Held as Part of HCI International 2009, San Diego, CA, USA, July 19-24, 2009. Proceedings 3*. Springer, 465–471.

[88] Lei Liu, Jean-Bernard Martens, and Robert Van Liere. 2011. Revisiting path steering for 3D manipulation tasks. *international Journal of Human-computer Studies* 69, 3 (2011), 170–181.

[89] Jacobus C Lock, Iain D Gilchrist, Iain D Gilchrist, Grzegorz Cielniak, and Nicola Bellotto. 2020. Experimental analysis of a spatialised audio interface for people with visual impairments. *ACM Transactions on Accessible Computing (TACCESS)* 13, 4 (2020), 1–21.

[90] Yujun Lu, BoYu Gao, Huawei Tu, Weiqi Luo, and HyungSeok Kim. 2021. Effects of Different Proximity-Based Feedback on Virtual Hand Pointing in Virtual Reality. In *Advances in Computer Graphics: 38th Computer Graphics International Conference, CGI 2021, Virtual Event, September 6–10, 2021, Proceedings 38*. Springer, 608–620.

[91] Paul Lubos, Gerd Bruder, and Frank Steinicke. 2014. Analysis of direct selection in head-mounted display environments. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 11–18.

[92] Francisco Lopez Luro and Veronica Sundstedt. 2019. A comparative study of eye tracking and hand controller for aiming tasks in virtual reality. In *Proceedings of the 11th ACM Symposium on eye tracking research & applications*. 1–9.

[93] Xiaolei Lv, Chengqi Xue, and Weiye Xiao. 2021. Influence of size and depth perception on ray-casting interaction in virtual reality. In *Intelligent Human Systems Integration 2021: Proceedings of the 4th International Conference on Intelligent Human Systems Integration (IHSI 2021): Integrating People and Intelligent Systems, February 22-24, 2021, Palermo, Italy*. Springer, 152–158.

[94] I Scott MacKenzie. 1992. Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction* 7, 1 (1992), 91–139.

[95] I Scott MacKenzie. 2018. Fitts' law. *The wiley handbook of human computer interaction* 1 (2018), 347–370.

[96] I Scott MacKenzie. 2024. *Human-computer interaction: An empirical research perspective*. Elsevier.

[97] I Scott MacKenzie and William Buxton. 1992. Extending Fitts' law to two-dimensional tasks. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 219–226.

[98] I Scott MacKenzie and Poika Isokoski. 2008. Fitts' throughput and the speed-accuracy tradeoff. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1633–1636.

[99] I Scott MacKenzie and Aleks Oniszczak. 1998. A comparison of three selection techniques for touchpads. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 336–343.

[100] I Scott MacKenzie and R William Soukoreff. 2003. Card, English, and Burr (1978) 25 years later. In *CHI'03 Extended Abstracts on Human Factors in Computing Systems*. 760–761.

[101] Daniel Medeiros, Graham Wilson, Mark Mcgill, and Stephen Anthony Brewster. 2023. The benefits of passive haptics and perceptual manipulation for extended reality interactions in constrained passenger spaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.

[102] Daniel Mendes, Fabio Marco Caputo, Andrea Giachetti, Alfredo Ferreira, and Joaquim Jorge. 2019. A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments. In *Computer graphics forum*, Vol. 38. Wiley Online Library, 21–45.

[103] David E Meyer, Richard A Abrams, Sylvan Kornblum, Charles E Wright, and JE Keith Smith. 1988. Optimality in human motor performance: ideal control of rapid aimed movements. *Psychological review* 95, 3 (1988), 340.

[104] Katsumi Minakata, John Paulin Hansen, I Scott MacKenzie, Per Bækgaard, and Vijay Rajanna. 2019. Pointing by gaze, head, and foot in a head-mounted display. In *Proceedings of the 11th ACM symposium on eye tracking research & applications*. 1–9.

[105] Mostafa Mohammadi, Hendrik Knoche, Michael Gaihede, Bo Bentsen, and Lotte NS Andreasen Struijk. 2019. A high-resolution tongue-based joystick to enable robot control for individuals with severe disabilities. In *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*. IEEE, 1043–1048.

[106] Pedro Monteiro, Guilherme Gonçalves, Bruno Peixoto, Miguel Melo, and Maximino Bessa. 2023. Evaluation of hands-free VR interaction methods during a Fitts' task: Efficiency and Effectiveness. *Ieee Access* (2023).

[107] Salim A Mouloua, Mustapha Mouloua, Daniel S McConnell, and PA Hancock. 2018. The effects of handedness and dominance on motor task performance. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 62. SAGE Publications Sage CA: Los Angeles, CA, 1237–1241.

[108] Atsuo Murata and Hirokazu Iwase. 2001. Extending Fitts' law to a three-dimensional pointing task. *Human movement science* 20, 6 (2001), 791–805.

[109] Jakob Nielsen. 1994. Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 152–158.

[110] Halla H Olafsdottir, Yves Guiard, Olivier Rioul, and Simon T Perrault. 2012. A new test of throughput invariance in Fitts' law: Role of the intercept and of Jensen's inequality. In *26th BCS Conference on Human Computer Interaction (HCI 2012)*.

[111] Matthew J Page, Joanne E McKenzie, Patrick M Bossuyt, Isabelle Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamseer, Jennifer M Tetzlaff, Elie A Akl, Sue E Brennan, et al. 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *bmj* 372 (2021), 9 pages.

[112] Yun Suen Pai, Tilman Dingler, and Kai Kunze. 2019. Assessing hands-free interactions for VR using eye gaze and electromyography. *Virtual Reality* 23 (2019), 119–131.

[113] Guillermo Palacios-Navarro, Sergio Albiol-Pérez, and Iván García-Magariño. 2015. A Kinect-Based Virtual Reality System for Parkinson Disease Rehabilitation. In *New Contributions in Information Systems and Technologies: Volume 1*. Springer, 1133–1139.

[114] Nikola Panic, Emanuele Leoncini, Giulio de Belvis, Walter Ricciardi, and Stefania Boccia. 2013. Evaluation of the endorsement of the preferred reporting items for systematic reviews and meta-analysis (PRISMA) statement on the quality of published systematic review and meta-analyses. *PloS one* 8, 12 (2013), e83138.

[115] Sai Akhil Penumudi, Veera Aneesh Kuppam, Jeong Ho Kim, and Jaejin Hwang. 2020. The effects of target location on musculoskeletal load, task performance, and subjective discomfort during virtual reality interactions. *Applied ergonomics* 84 (2020), 103010.

[116] Ivan Poupyrev, Suzanne Weghorst, Mark Billinghurst, and Tadao Ichikawa. 1998. A study of techniques for selecting and positioning objects in immersive VEs: effects of distance, size, and visual feedback. In *Proceedings of ACM CHI*, Vol. 98.

[117] John David Prieto Prada, Myung Ho Lee, and Cheol Song. 2023. A Gaze-Speech System in Mixed Reality for Human-Robot Interaction. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 7547–7553.

[118] Eric M Prager, Karen E Chambers, Joshua L Plotkin, David L McArthur, Anita E Bandrowski, Nidhi Bansal, Maryann E Martone, Hadley C Bergstrom, Anton Bespalov, and Chris Graf. 2019. Improving transparency and scientific rigor in academic publishing. , 377–390 pages.

[119] Kyle Reed, Michael Peshkin, J Edward Colgate, and James Patton. 2004. Initial studies in human-robot-human interaction: Fitts' law for two people. In *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004*, Vol. 3. IEEE, 2333–2338.

[120] Louis B Rosenberg. 1993. Virtual fixtures: Perceptual tools for telerobotic manipulation. In *Proceedings of IEEE virtual reality annual international symposium*. Ieee, 76–82.

[121] Greg S Ruthenbeck and Karen J Reynolds. 2015. Virtual reality for medical training: the state-of-the-art. *Journal of Simulation* 9, 1 (2015), 16–26.

[122] Anthony Scavarelli, Robert J Teather, and Ali Arya. 2023. Exploring Selection and Search Usability Across Desktop, Tablet, and Head-Mounted Display WebXR Platforms. In *2023 9th International Conference on Virtual Reality (ICVR)*. IEEE, 505–514.

[123] Joke Schuermans, Alena Van Hootegem, Margo Van den Bossche, Marjolein Van Gendt, Erik Witvrouw, and Evi Wezenbeek. 2022. Extended reality in musculoskeletal rehabilitation and injury prevention-A systematic review. *Physical therapy in sport* 55 (2022), 229–240.

[124] Gideon Schwarz. 1978. Estimating the dimension of a model. *The annals of statistics* (1978), 461–464.

[125] Valentin Schwind, Jan Leusmann, and Niels Henze. 2019. Understanding Visual-Haptic Integration of Avatar Hands Using a Fitts' Law Task in Virtual Reality. In *Proceedings of Mensch Und Computer 2019* (Hamburg, Germany) *(MuC '19)*. Association for Computing Machinery, New York, NY, USA, 211–222. https://doi.org/10.1145/3340764.3340769

[126] Jessica Sehrt, Tim Wißmann, Jan Breitenbach, and Valentin Schwind. 2023. The effects of body location and biosignal feedback modality on performance and workload using electromyography in virtual reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.

[127] Tawseef Ayoub Shaikh, Tabasum Rasool Dar, and Shabir Sofi. 2022. A data-centric artificial intelligent and extended reality technology in smart healthcare systems. *Social Network Analysis and Mining* 12, 1 (2022), 122.

[128] Masahiro Shiomi, Kayako Nakagawa, Kazuhiko Shinozawa, Reo Matsumura, Hiroshi Ishiguro, and Norihiro Hagita. 2017. Does a robot's touch encourage human effort? *International Journal of Social Robotics* 9 (2017), 5–15.

[129] Garth Shoemaker, Takayuki Tsukitani, Yoshifumi Kitamura, and Kellogg S Booth. 2012. Two-part models capture the impact of gain on pointing performance. *ACM Transactions on Computer-Human Interaction (TOCHI)* 19, 4 (2012), 1–34.

[130] Weixin Si, Xianavun Liao, Qiong Wang, and Phena-Ann Heng. 2018. Augmented reality-based personalized virtual operative anatomy for neurosurgical guidance and training. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 683–684.

[131] Shamus P Smith and Sam Du'Mont. 2009. Measuring the effect of gaming experience on virtual environment navigation tasks. In *2009 IEEE symposium on 3D user interfaces*. IEEE, 3–10.

[132] R William Soukoreff and I Scott MacKenzie. 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *International journal of human-computer studies* 61, 6 (2004), 751–789.

[133] Springer. 2004. Springer Link. https://link.springer.com. Last accessed: 2024-08-01.

[134] Kay M Stanney, Hannah Nye, Sam Haddad, Kelly S Hale, Christina K Padron, and Joseph V Cohn. 2021. Extended reality (XR) environments. *Handbook of human factors and ergonomics* (2021), 782–815.

[135] Anthony Steed and Chris Parker. 2005. Evaluating effectiveness of interaction techniques across immersive virtual environmental systems. *Presence* 14, 5 (2005), 511–527.

[136] Evropi Stefanidi, Marit Bentvelzen, Paweł W Woźniak, Thomas Kosch, Mikołaj P Woźniak, Thomas Mildner, Stefan Schneegass, Heiko Müller, and Jasmin Niess. 2023. Literature reviews in HCI: A review of reviews. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, New York, USA, 1–24.

[137] Frank Steinicke. 2016. *To Touch or Not to Touch?* Springer International Publishing, Cham, 107–124. https://doi.org/10.1007/978-3-319-43078-2_7

[138] Wolfgang Stuerzlinger and Robert J Teather. 2014. Considerations for targets in 3D pointing experiments. *Proceedings of HCI Korea* (2014), 162–168.

[139] Shilong Su, Juan He, Ruideng Wang, Zhengyang Chen, and Fang Zhou. 2024. The effectiveness of virtual reality, augmented reality, and mixed reality rehabilitation in total knee arthroplasty: A systematic review and Meta-Analysis. *The Journal of Arthroplasty* 39, 3 (2024), 582–590.

[140] Sriram Subramanian and WA IJsselsteijn. 2000. Survey and classification of spatial object manipulation techniques. In *Proceedings of OZCHI 2000-Conference on Human-Computer Interaction, Interfacing Reality in the New Millennium*. CSIRO, 330–337.

[141] Nizamettin T Tanacar, Moaaz H Mughrabi, Anil U Batmaz, Daniele Leonardis, and Mine Sarac. 2023. The impact of haptic feedback during sudden, rapid virtual interactions. In *2023 IEEE World Haptics Conference (WHC)*. IEEE, 64–70.

[142] Robert J Teather, Robert S Allison, and Wolfgang Stuerzlinger. 2009. Evaluating visual/motor co-location in fish-tank virtual reality. In *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*. IEEE, 624–629.

[143] Robert J Teather and Wolfgang Stuerzlinger. 2011. Pointing at 3D targets in a stereo head-tracked virtual environment. In *2011 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 87–94.

[144] Robert J Teather and Wolfgang Stuerzlinger. 2015. Factors affecting mouse-based 3d selection in desktop vr systems. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*. 10–19.

[145] Eleftherios Triantafyllidis and Zhibin Li. 2021. The challenges in modeling human performance in 3d space with fitts' law. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–9.

[146] AKM Amanat Ullah, William Delamare, and Khalad Hasan. 2023. Exploring Users Pointing Performance on Large Displays with Different Curvatures in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* (2023).

[147] Minna Vasarainen, Sami Paavola, and Liubov Vetoshkina. 2021. A systematic literature review on extended reality: Virtual, augmented and mixed reality in working life. *International Journal of Virtual Reality* 21, 2 (2021), 1–28.

[148] Radu-Daniel Vatavu and Ovidiu-Ciprian Ungurean. 2022. Understanding gesture input articulation with upper-body wearables for users with upper-body motor impairments. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–16.

[149] Jorge Wagner, Wolfgang Stuerzlinger, and Luciana Nedel. 2021. Comparing and combining virtual hand and virtual ray pointer interactions for data manipulation in immersive analytics. *IEEE transactions on visualization and computer graphics* 27, 5 (2021), 2513–2523.

[150] Uta Wagner, Mathias N Lystbæk, Pavel Manakhov, Jens Emil Sloth Grønbæk, Ken Pfeuffer, and Hans Gellersen. 2023. A fitts' law study of gaze-hand alignment for selection in 3d user interfaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–15.

[151] Samuel SY Wang. 2022. Use of Extended Reality in Medicine During the Covid-19 Pandemic. In *Extended Reality Usage During COVID 19 Pandemic*. Springer, 1–14.

[152] Benjamin Watson, Neff Walker, Peter Woytiuk, and William Ribarsky. 2003. Maintaining usability during 3D placement despite delay. In *IEEE Virtual Reality, 2003. Proceedings.* IEEE, 133–140.

[153] Hannah Weiss, Jianyang Tang, Connor Williams, and Leia Stirling. 2024. Performance on a target acquisition task differs between augmented reality and touch screen displays. *Applied Ergonomics* 116 (2024), 104185.

[154] Alan Traviss Welford. 1968. Fundamentals of skill. (1968).

[155] Tom Williams, Daniel Szafir, Tathagata Chakraborti, and Heni Ben Amor. 2018. Virtual, augmented, and mixed reality for human-robot interaction. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. 403–404.

[156] Dennis Wolf, John J Dudley, and Per Ola Kristensson. 2018. Performance envelopes of in-air direct and smartwatch indirect control for head-mounted augmented reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 347–354.

[157] Shota Yamanaka and Wolfgang Stuerzlinger. 2024. The Effect of Latency on Movement Time in Path-steering. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–19.

[158] Difeng Yu, Hai-Ning Liang, Xueshi Lu, Tianyu Zhang, and Wenge Xu. 2019. Depthmove: Leveraging head motions in the depth dimension to interact with virtual reality head-worn displays. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 103–114.

[159] Shumin Zhai. 2004. Characterizing computer input with Fitts' law parameters—the information and non-information aspects of pointing. *International Journal of Human-Computer Studies* 61, 6 (2004), 791–809.

[160] Shumin Zhai and Rogier Woltjer. 2003. Human movement performance in relation to path constraint-the law of steering in locomotion. In *IEEE Virtual Reality, 2003. Proceedings.* IEEE, 149–156.

[161] Lu Zhao, Yue Liu, Dejiang Ye, Zhuoluo Ma, and Weitao Song. 2020. Implementation and evaluation of touch-based interaction using electrovibration haptic feedback in virtual environments. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 239–247.

[162] Jingwen Zhi, Ruibiao Xie, Xiaowei Wang, Guotian Yang, Peng Peng, Xiao Luo, Yijun Chen, et al. 2020. A survey of Interactive Technology in Augmented Reality. In *2020 2nd International Conference on Information Technology and Computer Application (ITCA)*. IEEE, 101–104.

[163] Qian Zhou, George Fitzmaurice, and Fraser Anderson. 2022. In-depth mouse: Integrating desktop mouse into virtual reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.

[164] Xiaozhou Zhou, Yibing Guo, Lesong Jia, Yu Jin, Helu Li, and Chengqi Xue. 2023. A study of button size for virtual hand interaction in virtual environments based on clicking performance. *Multimedia Tools and Applications* 82, 10 (2023), 15903–15918.

[165] Yibo Zhu, Jagadish K Jayagopal, Ranjana K Mehta, Madhav Erraguntla, Joseph Nuamah, Anthony D McDonald, Heather Taylor, and Shuo-Hsiu Chang. 2020. Classifying major depressive disorder using fNIRS during motor rehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 28, 4 (2020), 961–969.

[166] Lukas Zimmerli, Carmen Krewer, Roger Gassert, Friedemann Müller, Robert Riener, and Lars Lünenburger. 2012. Validation of a mechanism to balance exercise difficulty in robot-assisted upper-extremity rehabilitation after stroke. *Journal of neuroengineering and rehabilitation* 9 (2012), 1–13.

[167] Sarah M Zweifach and Marc M Triola. 2019. Extended reality in medical education: driving adoption through provider-centered design. *Digital biomarkers* 3, 1 (2019), 14–21.

A Frequently used 3D XR Fitts' law equations

Table 7: Fitts' law equation variations for 3D XR.

Model Name	Interaction Space	3D display	Interaction Technique	Target Type	Studied Factor	Metric for Target Characteristic	Proposed Model
Shannon's [94]	2D Proximal	No	Direct	Planar Circle (Multidirectional)	A / W	Euclidean	$MT = a + b\log_2\left(\frac{A}{W}+1\right)$
Shannon's Effective [94]	2D Proximal	No	Direct	Planar Circle (Multidirectional)	effective A and W (A_e/W_e)	Euclidean	$MT = a + b\log_2\left(\frac{A_e}{W_e}+1\right)$
Angular [78]	2D distal	No	Extended	Vertical and Lateral	Angular Amplitude (α) and Width (ω)	Angular	$MT = a + b\log_2\left(\frac{\alpha}{\omega^k}+1\right)$
Angular Effective [16]	3D Both	Yes	Both	Spatial Circular (Multidirectional)	VAC* with Effective α and ω	Angular	$MT = a + b\log_2\left(\frac{\alpha_e}{\omega_e^k}+1\right)$
Murata and Iwase [108]	3D Proximal	No	Direct	Planar Circle (Multidirectional)	Azimuth* (φ)	Euclidean	$MT = a+b\,(\sin\varphi)+c\log_2\left(\frac{A}{W}+1\right)$
Cha and Myung [37]	3D Proximal	No	Direct	Spatial Circular (Multidirectional)	Azimuth* (φ) and Inclination* (θ)	Euclidean	$MT = a + b\varphi + c\sin(\theta) + d\cdot\log_2\left(\frac{2D}{W+F}\right)$
Machuca and Stuerzlinger [12]	3D Proximal	Yes	Direct	Depth Circular (Multidirectional)	Change in Target Depth (CTD*)	Euclidean	$MT = a+b\cdot\log_2\left(\frac{D_e}{W_e}+1\right)+c\cdot CTD$

This table summarizes key characteristics of Fitts' law variations in 3D and also 2D variations that are widely used in 3D XR studies. Note that target width and amplitude are used in all the provided models, as they are inspired by original Fitts' law [52]. VAC*: vergence-accomodation conflict [14]. **Azimuth*** or directional angle: direction of movement in a circular 2D layout, measured counterclockwise from the positive x-axis accounting for lateral movements [108]. **Inclination*** angle: vertical angle between the target and the positive y-axis in a spherical coordinate system to capture upwards and downward movements [37]. **CTD*** or Change of Target Depth: the depth change between consecutive target selections [12].

B Recommended Framework for Conducting 3D XR Fitts' Law Studies

Table 8: Recommended framework and practical implications

Checklist Elements	Research Theme (RT) and Recommendations
1. Study Objectives	
1.1. Study Goal and Research Theme	
Clearly define the goal within the related research theme (see subsection 5.2).	**GENERAL** - Recommendations applicable across all research themes. Explicitly defining the goal and theme of a study enhances transparency and provides a clearer distinction between types of studies. For instance, when examining the influence of a factor (EXP), specify whether the goal is to evaluate the effect of the factor using performance measures (e.g., movement time, error rate, throughput) or to propose a mathematical model (MDL). **Design (DES)** - The goal is to use Fitts' law to design interactions and interfaces. **Evaluation (EVAL)** - The goal is to evaluate a proposed system, interaction, or solution. **Exploration (EXP)** - The goal is to investigate the effect of external factors on 3D interaction, e.g., stereo-deficiencies or frame rate. **Modeling (MDL)** - The goal is to formulate new extensions to Fitts' law.
1.2. Hypotheses	
Formulate a hypothesis around quantitative and subjective measures (if applicable).	**GENERAL** - For quantitative measures, design a hypothesis around movement time (MT), error rate (ER), and throughput (TP), to provide insights on speed, accuracy, and speed-accuracy trade-offs. Use subjective measures (e.g., NASA-TLX [59], SUS [32]) for a broader understanding of perceived interaction quality (see subsection 5.4). **MDL** - Explicitly include the new factor under study along with the proposed model. Consider the effective Shannon form of Fitts' law, widely used in the research community, in both overall conditions and per-condition modeling for comprehensive evaluation and comparability (see subsection 5.1 and Table 7).
2. Methodology	
2.1. Model Selection	
Explicitly provide the selected Fitts' law extension and clear justification for the decision.	**GENERAL -** Each Fitts' law extension accounts for different factors and is validated in specific setups. We recommend selecting the extension that best aligns with the research goals and theme, as outlined in Table 7. For example, if the interaction technique is ray-based (not a direct one-to-one mapping), it is rational to choose the model that is empirically validated with a similar interaction technique that extends (Table 7) from the input device, instead of one-to-one mapping ("Direct" in Table 7). Additionally, since Shannon's model (Equation 6), included in ISO 9241-411 [67], is the most studied variation (see subsection 5.1), we recommend including it to ensure comparability with other studies. Reporting results based on Shannon's model also benefits the research community by facilitating insights across broader experimental setups and domains. Note that in Table 7, Murata and Iwase [108] and Cha and Myung [37] models are experimentally formulated to be on real-world interactions, not 3D XR setups (where various characteristics of 3D XR interaction such as vergence-accommodation [14, 23] and stereo deficiencies [12, 16] have been shown to influence user performance).
2.2. Task Design and Target Characteristics	
Report the shape of the target.	**GENERAL** - Clearly indicate the target shape in Fitts' task, specifying whether it is rectangular, circular, 2D (e.g., a disc), or 3D (e.g., a sphere or cylinder), as these shapes influence the perceived target area and hitpoint distribution [138]. Note that the sphere is widely used since it is the direct correspondence of ISO 9241-411 2D circles [67] within 3D space.

Continued on next page...

Checklist Elements	Research Theme (RT) and Recommendations
Report target arrangement in 3D space.	**GENERAL** - We recommend choosing the target arrangement to be the same as the one that was experimentally evaluated in the original model formulation the researchers decide on (Table 7). Note that the circular layout is the most studied variation, as also supported by ISO 9241-411 [67], and supports multidirectional selection (see subsection 5.5). Clearly indicate if consecutive selection with varying depths is allowed in one trial or if the task involves only the selection of 3D targets in a 2D planar layout with fixed depth at each trial (since moving in depth has been shown to be more challenging [12]). **MDL** - When designing a task for an experiment, the task should directly reflect and involve the specific factor being studied (see Table 7). **DES** - We recommend prioritizing the arrangement of the final system under study and refining common tasks accordingly. Note that the empirical evaluation of model fit in a refined task provides supporting insights regarding the validity and applicability of the selected model in the refined task. In this research theme, we recommend designing the task first to match the designed system and then to select the model based on the involved characteristics of interaction, e.g., interaction type (see Table 7).
Report target amplitude (A), width (W), and depth/distance (D) (see subsection 5.5).	**GENERAL** - Clearly indicate target W, A, and D in Fitts' tasks. Use Euclidean measures for consistency unless the selection extends beyond the peripersonal area, where alternative spatial metrics, i.e., angular measures, may be more appropriate since they provide measures based on the perceived values by the participant at a specific depth. We suggest avoiding the use of pixels (px) as units due to the low number of studies using them, especially when using Head-Mounted Displays (HMD), where resolutions vary widely.
Calculate and report Indices of Difficulty (ID).	**GENERAL -** Provide exact values of task difficulty used in the experiment as it can enhance the replicability of the user study and provide informative results, e.g., captured movement time, error rate, or throughput are reported for difficulties lower than 3 bits. The task difficulty recommended by ISO 9241-411 [67] for 2D interfaces is between 2-8. In the context of 3D XR interactions, we recommend covering IDs between 2-4.5 bits based on the observed trends in the reviewed studies (see subsection 5.4), which also increases the comparability of results. Following the 2D recommendation, we recommend using the adjustment of accuracy through ID_e [132]. Note that accuracy adjustment in Fitts' law accounts for spatial variability of human performance over a series of trials for calculating the main portion of hit points, i.e., effective width [95]. Besides using Equation 7, i.e., the "standard-deviation method", ID_e can also be calculated using the "discrete-error method" based on the error rate and $z-score$ [95]. Although the "discrete-error method" can be considered a simpler way of calculation due to its independence from capturing hit points and standard deviation, we recommend using the "standard-deviation method" since it relies on actual hit information from the experiment and provides more precise results [95], aligning with previous recommendations in 2D Fitts' studies [132]. **DES** - For XR user interface design, W_e can inform, e.g., the minimum size for targets to be reasonably easy to hit. Similarly, the relative task difficulty of a pointing task can be estimated through ID_e.
Vary directions for selection.	**GENERAL** - Including various directions for selection is recommended, e.g., left to right and right to left for lateral movements or varying directions in a circular arrangement (including clockwise and counterclockwise sequences based on the effect of handedness [12, 107] to ensure the generalizability of findings).

Continued on next page...

Checklist Elements	Research Theme (RT) and Recommendations
Report the selection strategy involved in the study and ask participants to act accordingly.	**GENERAL -** Since the selection strategy (as fast, as precise, or - most frequently - as fast and precise as possible) affects user performance, especially the movement time and error rate [19], we recommend guiding participants to select targets as 'fast and as precise as possible' to balance speed and accuracy in Fitts' studies, unless the study goal requires otherwise. Besides, based on the observed variability in studies, we recommend explicitly reporting if making an error is allowed in a trial that is considered a successful trial or not, to increase the transparency, applicability, and comparability of the user study. Note that Fitts' law originally considers single ballistic movements [52]. Thus, when errors are not permissible, multiple selection attempts may add substantial correction time to the movement and make movement time inconsistent across trials, ultimately making movement time not representative. Consequently, we do not recommend enforcing error-free trials for participants.
2.3. Interaction Space and Method	
Clearly explain the interaction technique incorporated in the study.	**GENERAL -** The interaction method is also useful for choosing the 3D variation of Fitts' law empirically validated in similar interaction (see Table 7), e.g., if the interaction is ray-based, 3D extensions originally validated with ray-based are recommended. Researchers can also choose other interaction techniques than the original method used, but providing the model-fit measures, e.g., R^2 is recommended to give insights on the validity and applicability of the used extension with incorporated interaction method. **DES/MDL -** For design and modeling we recommend the inclusion of ray casting and virtual hand interactions in Fitts' studies as they are the most common interaction methods and widely studied based on our findings (see subsection 5.3).
Report if any kind of 3D display is integrated into the study, or interaction is with real objects.	**GENERAL -** 3D interaction with real objects should be clearly distinguished from 3D interactions with virtual objects using stereo-displays or head-mounted displays, due to the known impact of various factors on user performance, e.g., the vergence-accommodation conflict [14].
Mention the system environment and interaction space used in the study.	**GENERAL -** Based on the findings of this paper, we recommend using head-mounted displays for more comparability of findings (see subsection 5.3) if the display is not determined by the study's aim. Using HMDs provides more comparability due to their wide usage in the investigated literature. Further, due to the lower number of studies compared to VR, investigating AR/MR/XR environments can also provide insights regarding the applicability and validity of Fitts' law extensions in XR. Also, a clear report of the environment used in the study potentially contributes to the replicability of the user study.
Explicitly mention the feedback type used in the user study, even if it is not an independent variable (see subsection 5.5).	**GENERAL -** As the type of feedback can affect user performance, we recommended it be reported to enhance the replicability and transparency of Fitts' studies. Highlighting targets (visual feedback) is recommended since it is the most common form of feedback in Fitts' studies based on the findings of this study, and it is also good user interface design practice [109]. Note that feedback type potentially influences user performance in Fitts' studies [19, 24, 101], so it should be reported as a part of the methodology. Regarding the visual feedback, we recommend choosing color vision deficiency-safe colors to ensure accessibility of the Fitts' study.
Participant selection and reporting.	**GENERAL -** Based on the findings regarding the participant numbers and demographics (see subsection 5.6), we recommend inviting at least 18 participants considering gender diversity and varying XR experience. Gaming experience should also be considered similarly, as increased gaming experience seems to affect 3D interaction [131, 144]
Choosing dependent variables (see subsection 5.4).	**GENERAL -** Based on the findings regarding the measurements used in investigated Fitts' studies, we recommend including movement time, error rate, and throughput to performance results regarding speed, accuracy, and speed-accuracy trade-off, respectively. Note that movement time is the time spent on the actual pointing movement, not other movements, i.e., reaction or dwell times should be excluded [132]. Regarding throughput, we recommend using ID/MT [94] instead of the slope formulation ($1/b$, also known as the index of performance) [132] due to the much larger number of studies using it, ensuring consistency and comparability. Further, we recommend the use of effective measures [67, 95], again to ensure comparability with other work.

Continued on next page...

Checklist Elements	**Research Theme (RT) and Recommendations**
3. Results, Analysis, and Report	
Plotting and visualizations.	**GENERAL -** We recommend using, at a very minimum, the observed step sizes in plots for movement time (0.5 s), error rate (5%), throughput (0.5 bits/s), and index of difficulty (0.5 bits) for easier visual comparison across different studies (see subsection 5.4). Particularly for the movement time versus index of difficulty plot, we recommend using seconds instead of ms, since the slope of the regression line presents the index of performance (bits/s), which enables better visual comparison of regression results.
Reporting dependent variables.	**GENERAL -** Clearly report captured movement time (s), error rate (%), and calculated throughput ($bits/s$) value as results, and do not solely rely on presentation through graphs, especially for throughput owing to the observed shorter variability change (see subsection 5.4). We recommend using the "means of means" method for throughput calculation ($TP = 1/n \sum_{i=1}^{n} \left(1/m \sum_{j=1}^{m} \left(ID_{ij}/MT_{ij}\right)\right)$, n and m are number of participants and conditions, respectively) instead of other alternative variations, e.g., average ID over average MT ($TP = \left(1/k \sum_{i=1}^{k} ID_k\right) / \left(1/k \sum_{i=1}^{k} MT_k\right)$, k is total number of trials). Note that each of these methods can lead to different results [110], making it important to highlight the variation used by researchers for improved comparability of findings. The "means of means" method can reduce the inter-subject variability impact on the calculated result, i.e., bias in cases where some participants provide higher contributions to throughput. Additionally, "means of means" has been previously recommended for 2D Fitts' studies [132, 159]. Mentioning the method for throughput calculation in the study also reduces inconsistencies in studies, improving the comparability and transparency of findings.
Movement time versus the index of difficulty regression.	**GENERAL -** One of the common visualizations of data in Fitts' experiments is the movement time and index of difficulty scatter plot along with the regression line representing the model used in the study [95]. However, just providing the regression line is not enough, and reporting the model-fit, e.g., R^2 value for showing the model-fit is required [132]. Note that for creating the scatter plots with linear regression, models are transformed into $MT = a + b \cdot ID$, meaning that more complex models will be simplified, e.g., Equation 10 is transformed into $MT = a + b' ID'$, where slope (b') is no longer the coefficient for Shannon's index of difficulty (Equation 6). This transformation needs to be reported clearly for transparency and replicability of the study. **MDL -** In comparing model-fits, due to varying numbers of coefficients present in complex Fitts' law 3D variations (see subsection 5.1 and Table 7) and the potential of overfitting while using R^2 [157], we recommend using adjusted R^2 instead. Further, to avoid artificial inflation of the quality of fit due to additional parameters, we recommend using metrics such as AIC [5] and/or BIC [124] to ensure outcomes that account for both the model's fit and number of free parameters of the equation itself, For example, for AIC it is typical to reject a model with $AIC > AIC_{min} + 10$ [34].
Optionally provide raw data for the experiment as supplementary material.	**GENERAL -** Given the increasing number of extensions of Fitts' law identified in this study (see subsection 5.1), researchers are encouraged to provide raw information on captured dependent variables for each combination of task design characteristics used in the study, e.g., studied factors in Table 7 (if applicable). This will enable future comparisons and collaborations by reusing the provided information in calculations with other models.

Continued on next page...

Checklist Elements	**Research Theme (RT) and Recommendations**
4. Post-Study Evaluation	
Trends and red flags for researchers and reviewers.	The observed trends from the reported user performance measures in Fitts' studies are useful to provide general insights for evaluating the captured results (see subsection 5.4). For example, movement times higher than 3 s and error rates above 40% for pointing tasks with difficulty less than 2 bits likely indicate a suboptimal interaction or a problem in the captured data that requires scrutiny (which in turn may point to problems in areas such as the logging, analysis, experimental design, or the underlying technology). Besides, a negative Index of Difficulty (ID) is also a red flag for evaluating the validity of findings. For example, in Shannon's Equation 6, which is shown in this study to be the most studied variation of Fitts' law, 0 bits is the minimum mathematically possible value for the ID ($log_2(1) = 0$). A negative ID can imply a serious theoretical issue with the model [95], and our findings indicate $ID > 0$ in all investigated empirical studies (see subsection 5.4). Based on the analysis of movement time versus ID regression slopes in investigated studies, a negative slope can also be considered a red flag. Besides, the conceptual meaning of a negative slope (b) in Fitts' studies would be a lower movement time in more difficult tasks, which (while spectacular) contradicts to current literature on human pointing.