# Comparing 3D Interaction Performance in Comfortable and Uncomfortable Regions

Marina Hofmann, Ronja Bürger, Ninja Frost, Julia Karremann, Jule Keller-Bacher, Stefanie Kraft, Gerd Bruder, Frank Steinicke

Immersive Media Group (IMG) Department of Computer Science & Department of Human-Computer-Media University of Würzburg E-Mail: {marina.hofmann,ronja.buerger,ninja.frost,julia.karremann, jule.keller-bacher,stefanie.b.kraft}@stud-mail.uni-wuerzburg.de, {gerd.bruder,frank.steinicke}@uni-wuerzburg.de

**Abstract:** Immersive virtual environments (IVEs) have the potential to afford natural interaction in the three-dimensional (3D) space around a user. While the available physical workspace can differ between IVEs, only a small region is located within arm's reach at any given moment. This interaction space is solely defined by the shape and posture of the user's body. Interaction performance in this space depends on a variety of ergonomics factors, the user's endurance, muscular strength, as well as fitness.

In this paper we investigate differences in selection task performance when users interact with their hands in a comfortable or uncomfortable region around their body. In a pilot study we identified comfortable and uncomfortable interaction regions for users who are standing upright. We conducted a Fitts' Law experiment to evaluate selection performance in these different regions over a duration of about thirty minutes. Although, we could not find any significant differences in interaction performance between the two regions, we observed a trend that the extent of physical fitness of the users affects performance: Athletic users perform better than unathletic users. We discuss implications for natural interaction in IVEs.

Keywords: Head-mounted displays, 3D interaction, comfortable interaction spaces

# **1** INTRODUCTION

Since the advent of 3D cinema and with 3D television and 3D gaming coming up in the consumer market, stereoscopic visualization is getting more and more important for a range of application fields. Moreover, advances in the field of unobtrusive body tracking, such as the Microsoft Kinect or the Leap Motion controller [lea13], afford natural interaction with 3D data sets. While stereoscopic display supports near-natural spatial impressions of virtual objects and scenes, using the hands and body to touch and manipulate virtual objects provides an intuitive direct interface for interaction with stereoscopically displayed 3D content. Immersive virtual environments (IVEs), such as tracked head-mounted displays (HMDs) or

CAVEs, thus have the potential to provide natural and intuitive interaction with virtual objects located in the proximity of the user's viewpoint. If an interactive virtual object is located within arm's reach, users can perform natural reach and touch gestures similar to the real world, whereas different IVEs provide users with different affordances for traveling to objects that are located at a larger distance. However, since an interactive virtual object could be located anywhere within arm's reach, the question arises how its position in different regions around the user's body affects comfort as well as interaction performance. Interaction in regions that users judge as comfortable, i.e., feeling physically relaxed without any pain or displeasing posture, could be beneficial for long interaction sessions. Performance may also differ depending on the muscle strength or flexibility of the user, and may be affected by the use of different muscle groups and levels of energy expenditure of the body. For instance, it is often observed that 3D direct interaction in mid-air as shown in different movies and TV series, such as *Minority Report* or in the *Iron Man* trilogy, costs significantly more musclular energy than desktop interaction, and may thus not provide high performance during prolonged use, although the performance during the first minutes may be encouraging.

In this paper we compare direct mid-air selection performance between users interacting in comfortable or uncomfortable regions within arm's reach. First, we performed a pre-study in which we determined regions that subjects rated as comfortable or uncomfortable. Then, we conducted a Fitts' Law experiment in which we compared 3D selection performance between groups interacting in these two regions. Our results show that subjects made fewer errors but required more time when interacting in a comfortable region. They made larger errors when interacting in an uncomfortable region. In contrast to our assumption, we found no significant main effect of the two conditions (uncomfortable and comfortable). However, we observed an effect of physical fitness on 3D selection performance.

In summary, our contributions include the

- analysis of comfortable and uncomfortable regions within arm's reach,
- comparison of selection performance in comfortable and uncomfortable regions, and
- guidelines for designing user interfaces with 3D direct mid-air interaction.

This paper is structured as follows. Section 2 gives an overview of related work. Section 3 explains the pilot study in which we identified comfortable and uncomfortable regions for 3D interaction. Section 4 describes the conducted Fitts' Law experiment. Results are shown in Section 5 and discussed in Section 6. Section 7 concludes the paper.

## 2 RELATED WORK

3D interaction in IVEs has been in the focus of many research groups over the last decades. Although direct interaction provides the most natural type of interaction with virtual objects, it is often not possible to use direct interaction for objects that are not located within arm's reach. Different indirect interaction techniques have been proposed, such as the Go-Go technique [PBWI96] and HOMER [BH97], which can provide users with the ability to interact with virtual objects in vista space by nonlinear scaling of hand positions within arm's reach. In particular, these techniques make use of the entire reachable space of a user's arms during interaction with distant objects, which may prove tiresome when constantly interacting at a distance, and may thus result in degraded performance over time. On the other hand, Mine et al. [MBS97] observed that such indirect interaction techniques result in degraded performance during interaction with virtual objects located within arm's reach. According to their results, direct interaction leads to significantly higher performance than manipulation of objects at a distance from the user's hand. Most results from similar studies agree on the point that optimal performance may be achieved when visual and motor spaces are superimposed or coupled closely [Dja98, LL07, WM99].

However, it is still an open research question, how the position of virtual target objects located within arm's reach may affect interaction performance. Direct interaction is subject to perceptual limitations, e.g., the vergence-accommodation mismatch, ghosting or double vision, which can result in significant misperception effects [BSS13a, BSS13b, CKC<sup>+</sup>10]. Depending on the location of virtual objects, users may be unable to discriminate interrelations or perceive distances to objects to be smaller or larger than they are displayed [LK03]. Distortions like that do not appear in real-world environments and may be related to limitations of current technology to correctly reproduce naturally occurring cues from the real world perfectly [WCCRT09]. Moreover, internal representations of space are influenced and updated by both visual and motoric input, which may affect interaction performance [WGTCR08, Tho83]. Due to varying energy expenditure between users based on differences in strength and endurance of arm muscles, interaction performance in mid-air within arm's reach in IVEs may be affected by different factors related to the ergonomics of direct interaction. In particular, contributing factors may include interaction duration, hand and arm postures, frequency of movements, and comfort.

**Fitts' Law** Fitts' Law describes the tradeoff between speed and accuracy in selection tasks [Fit54]. Selections by touching or grasping objects with a user's hands can be split up into two phases, the ballistic phase and the correction phase [LE08]. The ballistic phase consists of focusing on the target object and bringing the hand in the proximity of the goal by using proprioceptive motor control. After that, visual feedback is used in the correction phase in order to incrementally reduce the distance from the hand to the goal. Fitts' Law predicts the movement time MT for a given target distance D and size W. They are brought together in a log term which describes the difficulty of the task overall with  $MT = a + b \cdot log_2(D/W + 1)$ . The values a and b are empirically derived. The *index of difficulty* (ID) is given by the log term and indicates overall task difficulty; smaller or farther targets result in increased difficulty. The formula has been extended in order to get effective measures. The error rate is adjusted to 4% by resizing targets to their effective width  $W_e$ . This is supported by an international standard [Int00]. By calculating the average of the measured movement

distances,  $D_e$  can be determined. With that, the effective throughput can be computed as a useful combination of speed and accuracy:  $TP = \log_2 (D_e / W_e + 1) / MT$ . The validity of Fitts' Law for 3D interaction has been researched in the last years. Results from studies of several research groups imply that Fitts' Law is indeed valid for the kinematics of arm movements in a 3D interaction space [DKK07, MMD+87, MI08].

# 3 PILOT STUDY

While much research exists on ergonomics and comfort in Desktop environments, we found no previous research on comfort during 3D interaction in mid-air without passive haptic feedback as is common in IVEs. In this section we describe the pilot study that we conducted to determine comfortable regions in 3D mid-air interaction within arm's reach.

## 3.1 Study Design

In an informal study we recruited 27 participants (16 female, 11 male, age 14-57, M = 31.5) and asked them to take 26 predefined positions with their dominant hand and arm (see Figure 1). All subjects were righthanded. We defined the positions based on a grid with two different depths and six different heights. The different depths were realized by instructing the subjects to flex or to stretch out their arms.

In order not to bias the subjects, we provided them with minimal instructions as to the goals of the study. We randomized the positions and every subject had to hold every position for exactly 10 seconds. After every postion the subjects had to judge the comfort of that position using a 5-point ranking scale (1: very comfortable – 5: very uncomfortable). The total time per subject took less than 20 minutes.



Figure 1: Position grid as used in the pilot study.

#### 3.2 Results

We compared the means of the comfort values of every position (see Figure 2). The results imply that the most comfortable way to interact is close to the body with flexed arms. The most comfortable positions had a mean below 1: positions 27, 5, 6, 7, 10, and 11. In contrast, the least comfortable positions are the positions 3, 13, 24, 25, and 26. We observed



Figure 2: Means of the comfortable values of every position.

that positions with flexed arms are mostly considered more comfortable than positions with stretched-out arms. For the main study, the results were generalized into flexed arms at a distance of less than about 65% of the maximum arm's reach for comfortable positions and stretched-out arms at a distance of more than 65% of the maximum arm's reach for uncomfortable positions.

## 4 EXPERIMENT

In this section we describe the Fitts' Law experiment in which we analyzed touch behavior and performance in comfortable and uncomfortable regions.

## 4.1 Participants

In our experiment 10 male and 11 female subjects (ages 19-29, M = 21.9, heights 158-192 cm, M = 173.3 cm) had participated. All of the subjects were students of computer science or media communication and all of them received class credit for the participation in the experiment. Only right-handed subjects with normal or corrected to normal vision took part in the experiment. We measured the interpupillary distance (IPD) of each subject before the experiment using the technique proposed by Willemsen et al. [WGTCR08], which revealed IPDs between 5.5 cm and 7.0 cm (M = 6.3 cm). We used each individual's IPD for the experiment. At the beginning of the experiment, subjects were instructed to don the HMD and calibrate their virtual view. 21 subjects reported prior experience with stereoscopic 3D cinema (rating scale (0 = yes, 4 = no) M = 1.43), 8 subjects reported experience with HMDs (rating scale (0 = yes, 4 = no) M = 3.1), and 5 had previously participated in a study involving HMDs. The subjects were naive to the experimental conditions. The mean of the total time per subject was about 1 hour, with approximately 30 minutes immersed in the HMD.



Figure 3: Illustration of the experimental setup: (a) user wearing a tracked Sony HMZ-T1 HMD, and (b) virtual view on the HMD with target spheres arranged in a semicircle.

## 4.2 Materials

During the experiment subjects wore a Sony HMZ-T1 HMD. To track the user's head position and orientation for view-dependent rendering, we attached a passive 6-DOF target to the HMD and tracked it using an iotracker passive optical tracking system with 8 cameras. Also we fixed a 3-DOF target to the tip of the subject's index finger of the right hand to track its position and movements. The subject's left hand was placed on a Razer Nostromo keypad and used to confirm selections with fingertip movements of the right hand. The virtual stimulus used in the experiment consisted of a 3D scene, which was rendered with OpenGL on an Intel computer with Core i7 3.4GHz processors, 8GB of main memory and Nvidia Quadro 4000 graphics card. The targets in the experiment were represented by spheres. 7 spheres were arranged in a semicircle, with one sphere shown in the center, to select the sphere after every selection of a sphere from the circle (cf. [VBJS11]).

#### 4.3 Methods

At the beginning of the experiment, subjects were positioned standing in an upright posture in front of the tracking system (see Figure 3(a)). In an initial calibration phase we asked subjects to reach out as far as possible to the left, right, and straight down in front of body's center to compute the subject's interaction space within arm's reach. Then, subjects completed 3 to 10 supervised training trials for the experimental phase to ensure that they understood the task correctly. The training trials were excluded from the analysis.

We used a  $2 \times 13 \times 6$  design with the method of constant stimuli for the experiment trials. Interactions were performed either in a comfortable or uncomfortable region as suggested in Section 3, which was the only between-subjects variable in the experiment. The 6 radii and 13 heights of displayed target spheres were presented randomly and uniformly distributed between trials for each subject (see Figure 3(b)). Each trial consisted of sequential selections of all 7 targets in the semicircle with recurring selections of the one target in the center position, resulting in a total of 15 selections per trial. Subjects were instructed to select the targets as quickly and accurately as possible, as it is common in Fitts' Law experiments. After having selected a target correctly, subjects received feedback by targets turning green. We computed the distance of the index finger position to the calibrated sphere center, which indicated a selection if the distance was less than the sphere radius. If subjects performed a selection while the target sphere was not highlighted green, we recorded this as a selection error and advanced the trial state. The dependent variables were movement time, error distance (deviation from optimal target positions), error rate (percentage of targets missed), and effective throughput (see Section 2).

We used thirteen target heights (relative: from -0.6 to 0.6 in steps of 0.1), evenly split among the total vertical arm's reach of each participant. The absolute target height was centered on the elbow joint and scaled by the length of the forearm with the relative heights. We evaluated the 13 target heights with 6 radii scaled by the length of the forearm: comfortable interactions were representend with radii from 0.4 to 0.65 in steps of 0.05, uncomfortable interactions with radii from 0.7 to 0.95 in steps of 0.05. As discussed in Section 3, larger radii result in farther target distances, which correlate to less comfortable arm postures than short distances.

To be able to compare selection performance between targets displayed at different distances, we eliminated confounds of target distance on the results by scaling the size of target spheres according to Fitts' Law to a single *index of difficulty* of ID=3.25 (see Section 2). According to Fitts' Law, adapting the target size with respect to the distance between selections results in larger targets for longer selection distances, whereas the targets are smaller for shorter distances, thus resulting in the same task difficulty between the comfortable and uncomfortable interaction regions.

**Questionnaires** Additionally to the main experiment trials, we asked subjects to complete subjective questionnaires. Before and after the experiment subjects were asked to complete a Kennedy-Lane Simulator Sickness Questionnaire (SSQ). After the experimental phase subjects were asked to complete a Slater-Usoh-Steed (SUS) presence questionnaire.

## 5 RESULTS

In this section we summarize the results from the Fitts' Law experiment comparing the interaction performance in comfortable and uncomfortable regions. We had to exclude three subjects from the analysis due to simulator sickness symptoms; two of them performed interactions in uncomfortable regions. We analyzed the results with an unpaired two-sample t-test and a multiple univariate ANOVA at the 5% significance level.



Figure 4: Results of the Fitts' Law experiment: The plots show on the x-axis the trial and on the y-axis the (a) movement time, (b) error rate, (c) error distance, and (d) effective throughput for comfortable and uncomfortable interactions.

#### 5.1 Movement Time

The results for the movement time are illustrated in Figure 4(a). On average subjects required 1634.73 ms to move from one target to the next. We found no significant difference of movement time (t(16) = 0.61, p < .56) between comfortable (M = 1660.22 ms, SD = 195.37 ms) and uncomfortable (M = 1606.97 ms, SD = 187.01 ms) interactions.

## 5.2 Error Rate

The results for error rate are illustrated in Figure 4(b). The average error rate was 5.48 % (SD = 7.16 %). We found a trend for a difference of error rate t(16) = -1.56, p < .14) between comfortable (M = 4.46 %, SD = 2.09 %) and uncomfortable (M = 6.62 %, SD = 3.80 %) interaction.

#### 5.3 Error Distance

The results for the error distances between the calibrated center of each sphere and the finger position during selections are shown in Figure 4(c). The average error distance was 0.81 cm. We found a significant difference of error distance (t(16) = -8.36, p < .001) between comfortable (M = 0.77 cm, SD = 0.24 cm) and uncomfortable (M = 1.04 cm, SD = 0.47 cm) interaction. As expected from the Fitts' Law model, subjects interacting in the comfortable region worked more accurately than subjects interacting in the uncomfortable region due to the larger distances.

#### 5.4 Effective Throughput

The results for the effective throughput are illustrated in Figure 4(d). Throughput is a measure that incorporates both speed and accuracy. The higher the throughput the better. We found no significant main effect of effective throughput (t(16) = -0.20, p < .85) between the conditions. The average throughput during the experiment was M = 1.84 bps (SD = 0.21 bps) for subjects interacting in the comfortable region, while subjects in the uncomfortable region showed an average throughput of M = 1.86 bps (SD = 0.19 bps).

We investigated the effect of fitness and found a trend for a difference of effective throughput between physically fit and unfit subjects (t(16) = 1.89, p < .077). The average throughput during the experiment for fit subjects was M = 1.904 bps (SD = 0.146 bps), while unfit subjects had an average throughput of M=1.73 bps (SD=0.25 bps). We found no significant two-way interaction between fitness and condition (uncomfortable and comfortable) (F(1,14) = 0.32, p < .59).

We found a significant difference of throughput between male and female subjects (t(16) = -2.57, p < .05). The average throughput during the experiment was M = 1.735 bps (SD = 0.184 bps) for males, while females showed an average throughput of M = 1.95 bps (SD = 0.15 bps). However, we found no significant two way interaction between gender and condition (uncomfortable and comfortable) (F(1,14) = 0.001, p < .98).

#### 5.5 Questionnaires

Before and after the experiment, we asked subjects to judge their level of simulator sickness. On average we measured a pre-SSQ score of M = 0.11 (SD = 0.58) for comfortable and M = 0.16 (SD = 0.36) for uncomfortable interactions. The average post-SSQ scores was M = 0.49 (SD = 0.70) for comfortable and M = 0.5 (SD = 0.31) for uncomfortable interactions. The mean SUS-score for the reported sense of feeling present in the virtual scene was M = 2.54 (SD = 1.71) for comfortable and M = 2.94 (SD = 1.41) for uncomfortable interactions. We did not observe significant differences between the two conditions.

# 6 DISCUSSION

While we found a trend for error rate and a significant difference for error distance between interaction in the comfortable and uncomfortable regions, these results are in line with the predictions of the Fitts' Law model. In contrast to our expectations, we found no significant difference in effective throughput between comfortable and uncomfortable interaction. The results suggest that uncomfortable arm positions had no significant effect on 3D mid-air interaction performance over the course of the about 30 minutes of the experiment. As illustrated in Figure 4(d), the effective throughput did not appear to decrease over time. It remains an interesting question, whether interaction performance may decrease at a different speed in comfortable and uncomfortable regions over longer interaction sessions.

We also observed interesting between-subjects differences in 3D interaction performance. We found a trend for an increased interaction performance for physically fit subjects in our experiment. Also, female subjects had a significantly higher throughput than male subjects. These findings require further research.

We observed two main limitations of our Fitts' Law experiment, which may have affected the findings:

- Over the course of the experiment it became apparent that the results of the pilot study were generalized too much. In future research, a more distinct definition of (un)comfortable positions is needed.
- Furthermore, the combination of the weight and small 45 degrees field of view of the Sony HMZ-T1 HMD turned out to be a problem, as the subjects had to look downwards during the entire experiment. The resulting neck strain and uncomfortable body posture may have had a negative influence on the subjects' performance, in particular, in the "comfortable" regions. The larger field of view of current HMDs, such as the 110 degrees of the Oculus Rift, is likely to have a positive effect on the results.

In conclusion, despite the discussed problems, the results provide interesting vistas for future work on comfortable 3D interaction and behavior in IVEs.

# 7 CONCLUSION AND FUTURE WORK

Our presumption for the work described in this paper was that 3D interaction performance in IVEs depends on where the user interacts: in a comfortable or uncomfortable region within arm's reach. The results of our Fitts' Law experiment show no significant differences between the comfortable and uncomfortable regions that we have identified in a pilot study. However, we observed a trend that physically fit users gain a higher performance in 3D mid-air interaction than unfit users. While these findings help to understand how users interact with 3D stereoscopically displayed objects at different distances, further studies are required to fully analyze comfort during 3D interaction in IVEs. In particular, future work may encompass an investigation of endurance effects during 3D interaction by extending the duration of the experiment towards longer work sessions. Furthermore the scope of research can be extended by varying different parameters of the experiment, including the user's posture during interaction, different display systems, as well as tracking affordances such as provided by the Leap Motion controller.

# Literatur

[BH97] D. A. Bowman and L. F. Hodges. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In Proc. of ACM I3D, pages 35–38, 1997. [BSS13a] G. Bruder, F. Steinicke, and W. Stuerzlinger. Touching the void revisited: Analyses of touch behavior on and above tabletop surfaces. In Proc. of IN-TERACT, 17 pages, 2013. [BSS13b] G. Bruder, F. Steinicke, and W. Stuerzlinger. Effects of visual conflicts on 3D selection task performance in stereoscopic display environments. In Proc. of ACM 3DUI, pages 115–118, 2013.  $[CKC^+10]$ L.-W. Chan, H.-S. Kao, M. Y. Chen, M.-S. Lee, J. Hsu, and Y.-P. Hung. Touching the void: Direct-touch interaction for intangible displays. In Proc. of ACM CHI, pages 2625–2634, 2010. [Dja98] J. P. Djajadiningrat. Cubby: What You See is Where You Act. Interlacing the display and manipulation spaces. PhD thesis, Delft University of Technology, 1998. [DKK07] A. Y. Dvorkin, R. V. Kenyon, and E. A. Keshner. Reaching within a dynamic virtual environment. J. NeuroEng. Rehabil., 4(23):182-186, 2007. [Fit54] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. J. Exp. Psych., 47:381-391, 1954. [Int00] International Organization for Standardization. ISO/DIS 9241-9 Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9: Requirements for non-keyboard input devices, 2000. [LE08]G. Liu and J. T. Enns. Attention for perception and action: task interference for action planning, but not for online control. Exp. Brain Res., 185:709-717, 2008. [lea 13]Leap Motion. http://www.leapmotion.com, 2013.

- [LK03] J. M. Loomis and J. M. Knapp. Visual perception of egocentric distance in real and virtual environments. In Virtual and adaptive environments, pages 21-46. 2003.
- [LL07] D. K. Lemmerman and J. J. LaViola, Jr. Effects of interaction-display offset on user performance in surround screen virtual environments. In Proc. of IEEE VR, pages 303–304, 2007.
- [MBS97] M. R. Mine, F. P. Brooks, Jr., and C. H. Sequin. Moving objects in space: exploiting proprioception in virtual-environment interaction. In Proc. of ACM SIGGRAPH, pages 19–26, 1997.
- [MI08] I. S. MacKenzie and P. Isokoski. Fitts' throughput and the speed-accuracy tradeoff. In *Proc. of ACM CHI*, pages 1633–1636, 2008.
- [MMD+87] C. L. MacKenzie, R. G. Marteniuka, C. Dugasa, D. Liskea, and B. Eickmeiera. Three-dimensional movement trajectories in Fitts' task: Implications for control. Q.J. Exp. Psychology-A, 34(4):629-647, 1987.
- [PBWI96] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa. The Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR. In Proc. of ACM UIST, pages 79–80, 1996.
- [Tho83] J. A. Thomson. Is continuous visual monitoring necessary in visually guided locomotion? J. Exper. Psych. Hum. Percep. Perform., 9(3):427–443, 1983.
- [VBJS11] S. Vetter, J. Bützler, N. Jochems, and C. Schlick. Fitts' Law in bivariate pointing on large touch screens: Age-differentiated analysis of motion angle effects on movement times and error rates. In Proc. of Universal Access in Human-Computer Interaction. Users Diversity, International Conference (UAHCI), pages 620–628, 2011.
- [WCCRT09] P. Willemsen, M. B. Colton, S. H. Creem-Regehr, and W. B. Thompson. The effects of head-mounted display mechanical properties and field of view on distance judgements in virtual environments. ACM Trans. Appl. Percpt., 6(2):1– 14, 2009.
- [WGTCR08] P. Willemsen, A. A. Gooch, W. B. Thompson, and S. H. Creem-Regehr. Effects of stereo viewing conditions on distance perception in virtual environments. *Presence-Teleop. Virt.*, 17(1):91–101, 2008.
- [WM99] Y. Wang and C. MacKenzie. Effects of orientation disparity between haptic and graphic displays of objects in virtual environments. In Proc. of INTERACT, pages 391–398, 1999.