The Effect of Offset Correction and Cursor on Mid-Air Pointing in Real and Virtual Environments

Sven Mayer, Valentin Schwind, Robin Schweigert, Niels Henze
University of Stuttgart, Stuttgart, Germany, {firstname.lastname}@vis.uni-stuttgart.de

ABSTRACT
Pointing at remote objects to direct others' attention is a fundamental human ability. Previous work explored methods for remote pointing to select targets. Absolute pointing techniques that cast a ray from the user to a target are affected by humans' limited pointing accuracy. Recent work suggests that accuracy can be improved by compensating systematic offsets between targets a user aims at and rays cast from the user to the target. In this paper, we investigate mid-air pointing in the real world and virtual reality. Through a pointing study, we model the offsets to improve pointing accuracy and show that being in a virtual environment affects how users point at targets. In the second study, we validate the developed model and analyze the effect of compensating systematic offsets. We show that the provided model can significantly improve pointing accuracy when no cursor is provided. We further show that a cursor improves pointing accuracy but also increases the selection time.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords
Mid-air pointing; ray casting; modeling; offset correction; cursor; virtual environment.

INTRODUCTION
From early childhood on, humans have used mid-air pointing to direct others' attention [8]. Developing the skill to use referential gestures has been described as a pivotal change in infants' communicative competence and the foundation for engaging in conversations [8, 11]. Consequently, pointing plays an important role in human-computer interaction (HCI). Today's graphical user interfaces (GUIs) are essentially built around the user's ability to point at objects. Over the last decades, the effort went into building, evaluating, and refining pointing methods for GUIs to enable fast and precise input [57]. Today the input is mostly limited to mice, touchpads, and touchscreens.

CHI 2018, April 21–26, 2018, Montreal, QC, Canada

ISBN 978-1-4503-5620-6/18/04...$15.00

DOI: https://doi.org/10.1145/3173574.3174227

Beyond today's comment input devices, resent systems use the whole body as an input. Here we see mid-air pointing as one emerging input technique, and others have also been developed out of the early work by Bolt [10]. Plaumann et al. [53] inspired their investigation of mid-air pointing through smart environments such as smart home, while others facilitated mid-air pointing to interact with large high resolution displays (LHRDs) [38, 60]. Findings in the domain of LHRDs can also be adopted to improve the interaction with public displays, and other work such as Winkler et al. [61] used mid-air pointing to enrich the input space for a personal projector phone. Mid-air pointing has been proposed as one possible interaction technique for virtual content; for instance Argelaguet et al. [3] used mid-air pointing in a CAVE environment, using pointing as one collaborative tool to interact within a collaborative virtual environments [62]. Beyond simple mid-air pointing actions, a vast number of research projects investigated mid-air gesture sets e.g., [34, 37, 39].

Already Bolt's seminal work [10] demonstrated the potential of mid-air pointing to select remote targets. A large body of work investigated selecting remote physical and virtual targets. Previous work proposed relative and absolute input devices to enable remote pointing [7, 36, 42]. Early work was typically limited by the accuracy of the tracking technology. Absolute ray casting techniques enable users to use the same pointing gestures they use for communicating with other people but require tracking a user's hands or controllers with high precision. The recent revival of virtual reality (VR) has increased the need for fast and precise methods to point at objects in three dimensions. Current VR devices such as the HTC Vive and the Oculus Rift are delivered with controllers that enable a user to select virtual objects.

Although pointing in three dimensions to communicate with other humans is a fundamental human skill, work in experimental Psychology shows that humans' pointing accuracy is limited [19]. Recent work not only describes systematic errors when humans point at distant objects but also provides a first step towards modeling the error and compensating for systematic inaccuracies [40]. Mayer et al. [40] asked participants to point at crosshairs on a projection screen, measured the accuracy of different ray casting methods and provided a model to compensate the systematic offset for real-world (RW) mid-air pointing. While the work by Mayer et al. is promising and the authors conclude that they can improve pointing accuracy by 37.3%, the achieved accuracy is too low for precise selection, and the model has not been validated. Furthermore, it remains unclear if the model can be generalized to other contexts such as virtual reality and how it compares to a cursor that likely also improves pointing accuracy.

In this paper, we investigate the possible use of freehand mid-air pointing in the real and virtual environment. Further, we extend existing correction models and investigate the impact of visual feedback on humans' pointing performance. Therefore, we present two studies that investigate how humans point at targets in the real and virtual environment. In the first study, participants pointed at targets inside and outside VR. The results show that participants point differently while they are in VR. We argue that this is likely an effect caused by the VR glasses and the limited field of view. Using the collected data we developed models to compensate systematic offsets, which we validate in a second study. We show that the developed models can significantly improve pointing accuracy. We further show that a cursor can enhance mid-air pointing accuracy but thereby increases the selection time.

RELATED WORK

Previous work investigating mid-air pointing focused on the influences of psychology and physiology on pointing gestures, tracking techniques, mid-air ray cast techniques, offset compensation, and limb visualization in VR. In the following, we discuss these topics.

Psychology and Physiology

It has been shown that children in early childhood begin to express themselves with pointing gestures [25]. Pointing is linked to learning others' intentions and has a substantial impact on developing a theory of mind [12] as well as in associating verbal declarations [5]. Kendon [28] differentiates pointing gestures using the index finger, open hand, or thumb. While thumb and the open hand are used when the object being indicated is not primary focus or topic of the discourse, the extended index finger is used when a specific person, object, or location is meant [28]. Pointing requires a fine level of dexterity and motor control over intrinsic oscillations of the own body (tremor) as a result of involuntary, approximately rhythmic, and roughly sinusoidal movements [18]. Furthermore, both Christakos and Lal [13] and Riviere et al. [54] concluded that the hands move at $8\,to\,12\,Hz$ oscillations, and Basmajian and De Luca [6] stated that the oscillation is less than $13\,Hz$. Further, Morrison and Keogh [47] conducted a frequency analysis for pointing with the hand and index finger and found dominant frequency peaks between $2\,and\,4$ and between $8\,and\,12\,Hz$. They also found that oscillations increased when participants attempted to reduce the tremor by exerting greater control over the hand. Hand tremor was already described as an issue for HCI in an interaction scenario by Olsen and Nielsen [50] while using a laser pointer for selection tasks.

Ocular dominance is known to influence mid-air pointing [29]. Human ocular dominance can tested with, e.g., a test by Miles [43] and by Porta [16]. Plaumann et al. [53] confirmed these results using a high precision motion tracking system. Further, they concluded that handedness also has an influence on how humans point to distant targets.

Tracking Techniques

The user needs to be tracked to enable interaction from a distance. Related work presents two approaches for tracking. Either the user interacts with a controller or the user's body is tracked by surrounding equipment.

More and more computerized systems using a controller such as mice, keyboards or 3D input devices (e.g. Zhai et al. [63]) as the primary interaction device are now hitting the consumer market. In the domain of LHRDs most prototypes use controllers to overcome the distance between display and user [51]. We see the same trend in the game console market. Here even body movement focused game consoles like the Nintendo Wii [41], use a controller to recognize the body movement of the player. Even the latest technical innovation of augmented reality (AR) glasses, the Microsoft Hololense is shipped with a controller. Also VR glasses such as the Oculus Rift and the HTC Vive offer a controller for interaction with the VR scene. Third party technologies even provide the ability to track all ten fingers using gloves.

In contrast to controller and wearable systems passive systems can deliver the same richness of interaction without equipping the user. Nickel and Stiefelhagen [49] used RGB cameras with skin color tracking to approximate the pointing direction. While the LEAP Motion has been adopted to provide finger orientation to current VR glasses the overall detectable range is still limited. The limited range is mostly due to stereo vision reconstruction using two infrared cameras. To overcome the limited tracking possibilities most research prototypes simulate a perfect tracking using six-degree-of-freedom (6DOF) technologies, also known as motion capture systems. These passive tracking systems have widely been used over the last decade, for instance, by Kranstedt et al. [32] or Vogel and Balakrishnan [59, 60].

Mir-Air Ray Casting Techniques

In the following, we present absolute mid-air pointing ray casting techniques [60]. Mid-air pointing ray casting techniques can further be classified by the origin of the ray. Argelaguet et al. [3] distinguish between *eye-rooted* and *hand-rooted* techniques.

Two eye-rooted ray casting approaches are widely used; the eye orientation and the eye position as root of the ray. a) Using the eye orientation as a ray cast is refered to as gaze ray casting [49] and is implemented similar to pointing tasks using eye-tracking [35]. However, eye orientation ray casting requires special equipment and extra eye calibration. To avoid extra equipment and calibration, Nickel and Stiefelhagen [49] proposed using the orientation of the head; we refer to this technique as head ray cast (HRC). b) On the other hand are ray casting techniques which use the eye position as root of the ray. The most common technique eye-finger ray cast (EFRC), was specified in 1997 by Pierce et al. [52]. However, today EFRC, actually uses the "Cyclops Eye", which is the position between the eye, as root [32]. Kranstedt et al. [32] suggest that EFRC is defined by using the cyclops eye as root and the index fingertip as the direction.

Hand-root methods use the hand as the origin for the ray [45, 46]. Corradini and Cohen [15] identified index finger ray cast (IFRC) as the most common hand-rooted method. On the other hand, Nickel and Stiefelhagen [49] purposed and investigated an *elbow-rooted* ray casting method. We refer to this method as forearm ray cast (FRC).

Offset Compensation
Foley et al. [19] found a distance-dependent trend to overreach targets using pointing with the index finger. This finding was confirmed by Mayer et al. [40]. In their work, the authors describe systematic errors of absolute pointing and present a polynomial offset model for compensation. Akkil and Isokoski [1] conducted a study to compare different pointing techniques including eye gaze for compensation. Their results indicate that overlaying gaze information on an egocentric view increases the accuracy and confidence while pointing. On the other hand, Jota et al. [27] recommended using EFRC to reduce the parallax influence.

Visual Feedback
Wong and Gutwin [62] investigated different ways to visualize the pointing direction for VR. Their results suggest that a red line in the pointing direction is optimal for direction visualization. However, this is hard to realize in the RW. As a second option Wong and Gutwin [62] propose projecting a cursor on the object a user interacts with. In their implementation they used a red dot as cursor visualization. In an LHRD scenario Jiang et al. [26] used a red circle to visualize the cursors' position on a large display. Both "dot" and "circle" visualization can be realized in the RW using camera projector systems as provided by Benko et al. [9] and Gugenheimer et al. [23]. Kopper et al. [31] encoded the uncertainty of the position by mapping the amount of jitter to the circle size. Lastly, Nancel et al. [48] as well as Olsen et al. [50] used a red crosshair for their selection task. Furthermore, Cockburn et al. [14] investigated the effect of selection targets at a distance with and without visual feedback. They found that visual feedback improves selection accuracy. However, visual feedback might also influence the immersion in VR as Argelaguet and Andujar [2] showed that tracking technology, latency, and jitter influence the overall input performance.

Limb Visualization
As related work suggests using a finger and the forearm to indicate directions, it is necessary to visualize the arm and the hands to make mid-air pointing in VR feasible. Previous work found that the brain is able to accept virtual limbs [17] and bodies [56] as part of the own body. Rendering the own body in VR avoids fundamental limitations of human proprioception as the brain encodes limb positions primarily using vision [21, 22]. However, the illusion of body-ownership is affected by the visual appearance of the avatar. For example, Lin and Jörg [33] found that human-like hand models increased the illusion of body ownership and led to behavioral changes compared to more abstract representations. Similar findings were presented by Argelaguet et al. [4], who found that the appearance of avatars' hands in VR influences the user's sense of agency. However, the illusion of body ownership increases with human-like virtual hands. Schwind et al. [55] found a gender-related difference and, for example, recommended avoiding gender swapping in VR by using non-realistic or androgyny avatars. Furthermore, research comparing input methods in VR and real world found that VR is still limited. For example, Knierim et al. [30] compared the typing performance of users in the real and virtual world. Their results show that the typing performance of users in the virtual world is limited and depends on their experience of the users.

Summary
A substantial body of research has investigated the selection of distant targets. Previous work has shown that interaction without a controller is hard to implement, however it has also been shown that carrying no controller has its advantages. In this paper, we focus only on absolute mid-air pointing without using a controller. Mayer et al. [40] presented a systematic offset between the ray cast and the target for the RW. However, they have not tested how effective the model is in a real selection task. Further, the model has not been applied to a real selection task, thus the impact on task completion time (TCT) is unknown. Due to the rise of AR and VR availability it also would be interesting to see how the model performs in different environments.

To address these open questions, we replicate the work by Mayer et al. and extend it by also determining offset models for VR. We then apply the models in a real selection task to ensure the external validity of the developed models. Since previous work did not apply and validate their model, we investigate how the model performs in RW and VR regarding offset and TCT. Further, as related work suggested using a cursor for precise input, we investigate the effect of displaying a cursor and how a cursor affects offset and TCT.

DATA COLLECTION STUDY
We conducted the first study to record labeled body postures while participants performed mid-air pointing gestures. Our goal was to compare RW and VR. Thus, participants were asked to perform mid-air pointing gestures in both environments. Differences between the two environments would suggest that to correct the systematic error, separate models are needed. As Mayer et al. [40] showed that angular models are sufficient for all pointing distances, we only investigate standing in $2\,m$ distance to the target. Further, as presenting feedback might change the users behavior we did not present any feedback to the user to record natural mid-air pointing gestures. Moreover, to build a precise model we needed to present targets without an area. This is in line with Mayer et al. [40]. No target area means that the target becomes a single coordinate on the projection canvas. This allowed us to build a model without possible misinterpretation by participants, as pointing on a target with an area might convey the message of pointing onto the center or somewhere on the target area.

Study Design
We used a within-subject design with a single independent variable (IV): ENVIRONMENT. The IV ENVIRONMENT has two levels: *RealWorld* and *VirtualReality*. We replicated the setup of Mayer et al. [40], and also used 35 targets in a 7×5

Figure 1. One participant pointing at a target in RW.

grid. Participants had to point 6 times on each target per condition resulting in a total of 420 pointing tasks. The order of the targets was randomized while the order of ENVIRONMENT was counter-balanced.

Apparatus

As apparatus, we used a PC running Windows 10 connected to a projector, a head-mounted display (HMD), and a marker-based 6DOF motion capture system namely an OptiTrack system. As HMD we used an HTC Vive. To guarantee a smoothly running VR experience we used a NVIDIA GeForce GTX 1080. The tracking system delivers the absolute position of the markers attached to the participant at $30\,FPS$. We calibrated the system as suggested by the manufacturer resulting in millimeter accuracy. The software to interact with the tracking system provides a full-body tracking by attaching a number of markers. However, as the software is closed source and approximates the position of body parts, especially the fingertip, we did not use OptiTrack's commercial full-body tracking implementation. Instead, we used 7 rigid bodies to track the body without any approximations. We tracked the head/HMD, both shoulders, the right upper arm, the right lower arm, the hand root, and the index finger as shown in Figure 2. We used markers with a diameter of $15.9\,mm$ and $19.\,mm$ to ensure a stable tacking. We 3D printed custom mounts[1] to determine the pose of the right arm, the hand, the index finger, and the HTC Vive. As depicted in Figure 2 the index finger marker is wrapped around the finger and the upper as well as the forearm marker are wrapped around the arm.

To perfectly represent the participant in VR we took the following measurements: We took the precise length of the index finger, hand, lower and upper arm and measured the diameter of the finger, hand, wrist, elbow, lower arm, upper arm and head. Further, we took measurements of both shoulder and eye position in relation to the reflective shoulder markers. Lastly we measured the position of the lower and upper arm markers in relation to the elbow. The tracked positions of the marker combined with these 14 participant specific measurements enabled us to precisely determine the position and orientation of the upper body, the arm, and the index finger. We adjusted the avatar's dimensions as well as the underlying bone structure to precisely represent the participant's real dimensions.

These measurements also have been used to calculate the perfect ray-casts for all 4 mid-air ray casting techniques:

Index finger ray cast (IFRC): Using the finger tip marker plus a user-specific marker placement measurement we calculate the true finger tip position. Additionally we used the finger tip markers orientation to determine the direction of the ray.

Head ray cast (HRC): We used the *Cyclops Eye* ray cast as proposed by Kranstedt et al. [32]. Therefore, in the VR condition, we used the HMDs maskers to calculate the position of the bridge of the nose and its forward direction. On the other hand, in the RW condition, we used a marker on the head of the participant plus head measurements to also determine the bridge of the nose and the forward direction of the head.

Eye-finger ray cast (EFRC): The root for the ray cast, *Cyclops Eye* calculated the same way for the HRC. The finger tips position was optioned in the same way as for the IFRC and used as the direction vector.

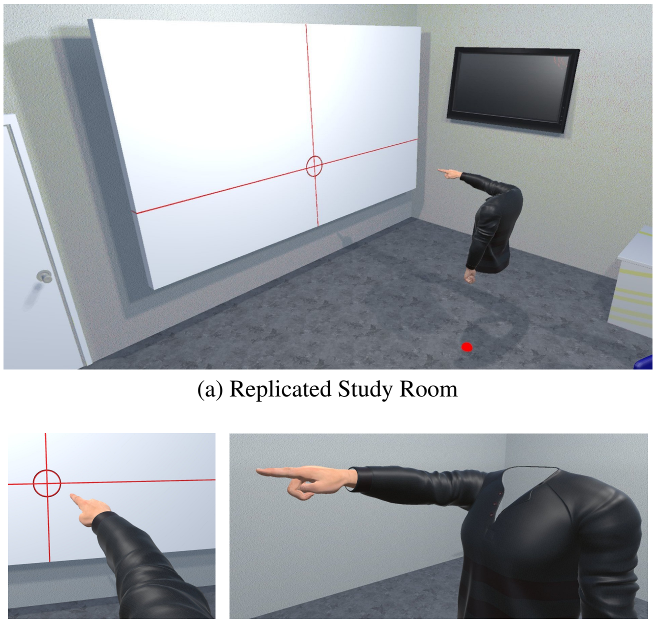
Forearm ray cast (FRC): We calculated the center or the forearm by approximating the forearm with a frustum of a cone. This was achieved using the position and orientation of the forearm marker plus additional measurements.

The 35 presented targets were arranged in a 7×5 (column $\times$ row) grid. The targets were either projected on a projection screen ($269.4\,cm \times 136.2\,m$) or presented in VR on the same sized the virtual projection screen. The spacing of the target grid was $44.9\,cm \times 34.\,cm$.

Both VR scene and RW projection were implemented using Unity version 5.6. The projector mounted in the study room projected the targets rendered in the VR scene to avoid alignments issues. We therefore designed the VR scene to replicate the real room the participants were standing in, see Figure 3a. To ensure a precise representation of the room in VR we used a professional laser measurement tool (accuracy $\pm 1.5\,mm$). We recreated the room in VR to avoid any interference on the pointing performance and to keep the results comparable. As humans use their hand as a reference point for mid-air pointing, it is important to represent them accurately in VR.

Figure 2. The seven rigid body markers used for body tracking in our study.

[1] 3D models of the custom mounts used in our study: `github.com/interactionlab/htc-vive-marker-mount`



(a) Replicated Study Room

(b) Hand (c) Hand and Body

Figure 3. The VR scene we used in our study.

Therefore, we used the additional 14 participant specific measurements to ensure a precise visualization of the user's arm and hand. Furthermore, Schwind et al. [55] showed an effect of hand representation on the feeling of eeriness of the participants. Thus, we used the same human androgynous hands[2] as these caused the smallest gender-related effect on participants' acceptance. The hand representation is shown in Figures 3b and 3c.

Procedure

We followed the instructions and procedure that Mayer et al. [40] used to record their ground truth data. After welcoming a participant, we explained the procedure of the study and asked them to fill an informed consent as well as a demographic questionnaire. Afterward, we took 14 participant specific measurements to have a perfect representation of the arm and hand in VR. We asked them to stand at a specific position in the room (in RW the point was marked on the floor and in VR the point was indicated by a red dot on the floor, see Figure 3a) which was centered $2\,m$ away from the projection screen. From this point, participants were asked to aim at the targets using their dominant hand.

To compensate for natural hand tremor, described as an issue by Olsen and Nielsen [50], participants had to hold the pointing position for one second. To ensure this time span, participants had to click with the non-dominant hand on the button of a remote control when they started to hold a gesture. The target disappeared after one second. We instructed the participant to point as they would naturally do in other situations. We intentionally did not restrict participants body pose to record a range of pointing postures. In total the participates had to perform 420 mid-air pointing gestures. We split these into 4 sessions each with 105 gestures. Between the sessions, we asked them to fill a raw NASA-Task Load Index (raw TLX) [24] to check for fatigue effects. We randomized the target order and counter-balanced the order of ENVIRONMENT.

[2]Source for human androgynous hands we used in out study: `github.com/valentin-schwind/selfpresence`

Participants

We recruited participants from our university's volunteer pool. In total, 20 participants took part in the study (4 female, 16 male). The age of the participant was between 17 and 30 ($M = 22.1$, $SD = 3.1$). The body height was between $156\,cm$ and $190\,cm$ ($M = 175.2$, $SD = 9.9$). As Plaumann et al. [53] showed a strong influence of handedness we only recruited right-handed participants who had no locomotor coordination problems. We used the Miles [43] and Porta test [16] to screen participants for eye-dominance. 10 participants had right-eye dominance, 6 had left-eye dominance, and 4 were unclear.

Results

We collected a total amount of 8,400 mid-air pointing postures. For all of them, we calculated the following four different ray casting methods (METHOD): eye-finger ray cast (EFRC), index finger ray cast (IFRC), forearm ray cast (FRC), and head ray cast (HRC).

Fatigue effect

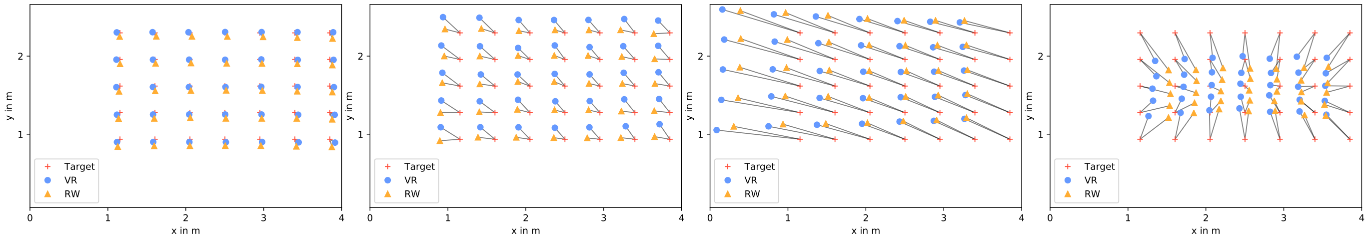
First, we analyzed the raw TLX score to determine if potential workload or fatigue effects had to be considered in the further analysis. The mean raw TLX score was $M = 35.42$ ($SD = 10.46$) after the first, $M = 35.38$ ($SD = 12.31$) after the second, $M = 35.46$ ($SD = 15.37$) after the third, and $M = 36.$ ($SD = 16.15$) after the last session. We conducted a one-way repeated measures analysis of variance (RM-ANOVA). As the analysis did not reveal a significant effect, $F_{3,57} = .047, p = .986$, we assume that the effect of participants' fatigue or workload was negligible.

Preprocessing

To determine the cast rays for each mid-air pointing postures, we used the samples between $100\,ms$ and $900\,ms$ to counteract possible hand tremor and possible movements at the beginning and end of the pointing phase. We further defined the offset as the distance between the position where the ray cast intersects with the projection screen and the position of the target. We then filtered the mid-air pointing postures to remove outliers using two times the standard deviation as an upper bound. Related work has shown that the head is the origin of human pointing. However, the participants were of different sizes so to compensate for different heights we aligned the heads of the participants to build one universal model.

Accuracy of Ray Casts

Table 1 shows the average offsets for ENVIRONMENT and METHOD respectively. The average offset is $9.33\,cm$ for EFRC, $28.09\,cm$ for IFRC, $65.\,cm$ for FRC and $42.46\,cm$ for HRC. We performed four one-way RM-ANOVAs to determine if the variance within one ray casting method is different in the *RealWorld* compared to the *VirtualReality*. We found a statistically significant difference for EFRC, $F_{1,19} = 5.845$, $p = .026$, FRC, $F_{1,19} = 33.13$, $p < .001$, and HRC, $F_{1,19} = 31.48$, $p < .001$. However, we found no statistically significant difference for IFRC, $F_{1,19} = .447$, $p = .512$.



(a) Eye-finger ray cast (EFRC) (b) Index finger ray cast (IFRC) (c) Forearm ray cast (FRC) (d) Head ray cast (HRC)

Figure 4. The average offset for each of the four ray cast techniques.

Environment	EFRC		IFRC		FRC		HRC	
	M	SD	M	SD	M	SD	M	SD
Distance *RealWorld*	10.21	5.56	29.14	19.24	60.34	20.16	47.15	23.71
Distance *VirtualReality*	8.45	4.54	27.03	18.95	69.66	25.42	37.77	19.19
Distance after correction *RealWorld*	8.02	4.54	15.17	9.15	27.01	12.36	46.96	23.71
Distance after correction *VirtualReality*	8.41	4.56	12.90	7.94	30.18	13.83	37.81	19.26

Table 1. Overall offsets between interact and target. Distance are reported in cm.

Modeling

As Mayer et al. [40] we built models to compensate the systematic error. Therefore we first define α_{pitch} as the vertical deviation angle and α_{yaw} as the horizontal deviation angle each between the ray cast and the body. Further Δ_{pitch} and Δ_{yaw} are the two correction angles respectively.

We used the following four functions also used by Mayer et al. [40]. The first function $f_1(\alpha_\omega)$ is a one-dimensional second degree polynomial function (parabola) to predict the correction Δ_ω. For $_\omega$ we use $_{pitch}$ or $_{yaw}$ to predict Δ_{pitch} and Δ_{yaw}. For the rest of the models, we are using α_{pitch} and α_{yaw} to predict Δ_ω. The functions $f_2(\alpha_{pitch}, \alpha_{yaw})$ and $f_3(\alpha_{pitch}, \alpha_{yaw})$ are complete two-dimensional polynomial functions, where f_2 is of degree 1 and f_3 of degree 2. The function $f_4(\alpha_{pitch}, \alpha_{yaw})$ is the function which performed best for Mayer et al. [40] to compensate the offset:

$$\begin{aligned} f_4(\alpha_p, \alpha_y) = {} & x_{14}\alpha_p^4 + x_{13}\alpha_y^4 + x_{12}\alpha_p^3\alpha_y + x_{11}\alpha_p\alpha_y^3 + \\ & x_{10}\alpha_p^3 + x_9\alpha_y^3 + x_8\alpha_p^2\alpha_y^2 + x_7\alpha_p^2\alpha_y + x_6\alpha_p\alpha_y^2 + \\ & x_5\alpha_p^2 + x_4\alpha_y^2 + x_3\alpha_p\alpha_y + x_2\alpha_p + x_1\alpha_y + x_0 \end{aligned} \tag{1}$$

While x_0 to x_{14} are the 15 parameters to fit. We used a nonlinear least-squares solver to fit out data.

Since we found that three ray cast METHODS are significantly different for *RealWorld* and *VirtualReality*, we fit models independently for each ENVIRONMENT. For a first evaluation of the models, we used leave-one-out cross-validation (LOOCV). We found that f_4 performed best with an overall correction of 29.3%. We achieved the best correction with FRC (55.9%) than IFRC with 50.1% then EFRC with 10.9% then HRC with .2%. However, the remaining offset was the smallest with EFRC (8.2 cm) then IFRC with 14. cm then FRC with 28.6 cm and the biggest when using HRC with a remaining error of 42.3 cm. The average improvement results using LOOCV are reported in Table 1.

Model Discussion

We found statistically significant differences between *RealWorld* and *VirtualReality* for EFRC, FRC, and HRC but not for IFRC. We assume this is due to the limited field of view (FoV) of the HMD. As depicted in Figure 4d VR caused more head movement than RW. More head movements reduce the offset between head ray and actual target, resulting in a lower HRC offset for VR. Furthermore, this also reduces the offset for EFRC as here the head is used to calculate the cyclops eye for the ray. The already reduced offset limits the possibility for an offset correction. Thus we only achieved a reduction of .5% for the VR EFRC model.

As our new model fits best using a two-dimensional polynomial and fits the best for offset correction for the RW, we confirmed the offset correction model presented by Mayer et al. [40] in the RW. We also showed that the same polynomial also reduced the offset the best for VR even though we found a significant difference between *RealWorld* and *VirtualReality*. However, we could not confirm that IFRC outperforms EFRC the remaining error after correction for RW. We found that the offset for IFRC is 89.2 % larger than EFRC while Mayer et al. [40] found that EFRC is 4.9 % larger than IFRC (for 2 m standing). However, before correction, they also reported that EFRC outperforms IFRC.

Overall Mayer et al. [40] reported errors before correction 4.8 times larger for EFRC, 1.9 for IFRC and 3.7 for FRC than the errors of the presented study. We believe this is due to their different tracking method. While Mayer et al. used one marker for each position and a post process labeling step, we used at least three markers per position of interest (e.g. fingertip). This enabled us to monitor participants' movements in real time which was necessary for the VR visualization, and also contributed towards a more stable and precise tracking.

While the offsets reported by Mayer et al. [40] are larger than the offsets we found, the overall direction is the same. They reported that the intersect is shifted to the upper left for IFRC and FRC while EFRC is shifted to the lower right. As depicted in Figure 5 we can confirm these findings for VR as well as RW. As we also investigated HRC here, we see a different trend. The offsets are shifted towards the center of the grid. Our HRC method is only derived from the head movement. Thus the eye movements are neglected in our implementation. The difference of the eye ray and the head ray could explain the effect of a shift towards the center as participants always focus on the target with their eyes. This can be confirmed with findings from the field of neurophysiology which studied the coordination of eye, head, and body movements in detail. Here, John S. Stahl [58] found that "head movements are orchestrated to control eye eccentricity". Further, Freedman and Sparks [20] found that humans even rotate their head to focus on the target while at the same time minimizing the effort put on ocular muscles. However, another factor could again be the limited FoV of the HMD.

EVALUATION

To validate the developed models and investigate the effect on users' performance we conducted a second study. As eye-finger ray cast (EFRC) resulted in the lowest offset, we tested the effect offset correction on participants' performance using EFRC. We were interested in testing the models in the real world as well as in VR. In contrast to our first study, we also investigated how the model performs when visual feedback is presented to the participant. Again we used targets without a target size to evaluate the models' performance. We used ENVIRONMENT (with levels *RealWorld* and *VirtualWorld*), CORRECTION (*Yes* and *No*), and CURSOR (*Yes* and *No*) as IVs. As dependent variables (DVs) we measured pointing precision, the TCT, and again used raw TLX questionnaires. We use the distance between a target's center and the intersection of the ray cast with the projection screen as accuracy. TCT is the time between the appearance of the target and the selection by the participants, as confirmed by a button on a remote control pressed with the non-dominant hand.

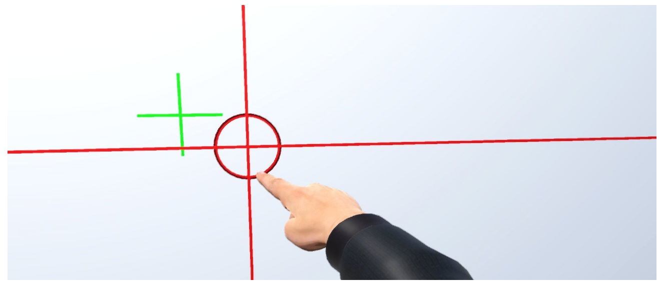
Study Design

We employed a $2 \times 2 \times 2$ factorial design for the second study. However, the conditions *No* CURSOR with or without correction were the same for the participant for both *RealWorld* and *VirtualWorld*, so the correction could not be noticed by the participant during the study. Therefore we were able to reduce the number of conditions to 6 while internally applying the correction or not to get all 8 conditions. With 2 repetitions per condition, we managed to keep the trials manageable for the participant and the time reasonable. Thus we had $6\,conditions \times 35\,targets \times 2\,respiration = 420\,trails$, which the participants completed in approximately one hour.

Apparatus

The overall setup was the same as in the first study. We used the same tracking system, optical markers, 35 targets, HMD, projector, and software. However, the Unity scene was adjusted to support our model if needed as well to support the visual feedback CURSOR. The visual feedback CURSOR was represented by a green crosshair as suggested by Olsen and Nielsen [50].

(a) The RW scene.



(b) The VR scene.

Figure 5. The RW and VR scenes used in our evaluation study while the green cursor is visible.

Procedure

After welcoming a participant, we explained the procedure of the study and asked him/her to fill an informed consent as well as a demographic questionnaire. Afterward, we took 14 measurements of the participant to have a perfect representation in VR. Participants had to press the button of a remote control with their non-dominant hand when they were confident that they wanted to continue with the next target. However, after the button press, we implemented a random $.5\,sec$ to $1.\,sec$ delay to ensure that participants did not move their arm before the next target appeared, to counteract possible false starts. As in the first study, we asked participants to stand at a specific position in the room centered $2\,m$ away from the projection screen and point at the targets using their dominant hand. We further instructed them to point as they would naturally do in other situations, but as quickly and accurately as possible. We intentionally did not restrict their body pose to record a range of pointing postures. After each condition we let participants fill a raw TLX questionnaire. All targets were randomized. CORRECTION and CURSOR were randomized within ENVIRONMENT while ENVIRONMENT was counter-balanced.

Participants

We recruited new participants from our university's self-volunteer pool. In total, 16 participants took part in the study (1 female, 15 male), aged between 19 and 26 ($M = 22.7$, $SD = 1.8$). The body height was between $156\,cm$ and $181\,cm$ ($M = 170.4$, $SD = 7.1$). All of them were right-handed, and none had locomotor coordination problems. We again used the Miles [43] and Porta test [16] to screen participants for eye-dominance. Ten had right-eye dominance, 1 left-eye dominance, and 5 were unclear.

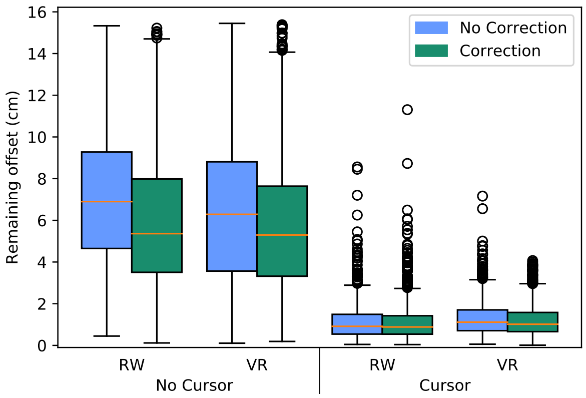


Figure 6. Remaining offset between interact and target for CORRECTION × CURSOR × ENVIRONMENT.

Results

In the following, we present the results of our correction models applied on eye-finger ray cast (EFRC) for *RealWorld* and *VirtualWorld*. We conducted a three-way RM-ANOVA with the independent within-subject variables CURSOR (with the levels *Yes* and *No*) vs. CORRECTION (*Yes* and *No*) vs. ENVIRONMENT (*RealWorld* and *VirtualWorld*). Since all factors had only two levels, no pairwise post-hoc comparisons were conducted. We used the distance between the ray cast using eye-finger ray cast (EFRC) and the target as accuracy measure and TCT as an indicator of the participants' pointing performance.

Fatigue effect

First, we again analyzed the raw NASA-Task Load Index (raw TLX) score to determine if potential workload or fatigue effects had to be considered in the further analysis. The mean raw TLX score was $M = 36.25$ ($SD = 11.37$) after the first, $M = 38.28$ ($SD = 13.46$) after the second, $M = 39.84$ ($SD = 15.10$) after the third, $M = 37.81$ ($SD = 16.32$) after the fourth, $M = 39.74$ ($SD = 15.73$) after the fifth, $M = 37.76$ ($SD = 17.68$) after the last session. We conducted a one-way RM-ANOVA. As the analysis did not reveal a significant effect, $F_{5,15} = .654, p = .659$, we again assume that the effect of participants' fatigue or workload was negligible.

Accuracy

We found a significant effect of CORRECTION, $F_{1,15} = 5.321, p = .027$, CURSOR, $F_{1,15} = 131.9, p < .001$, and ENVIRONMENT, $F_{1,15} = 1.3, p = .027$ on the participants' pointing accuracy. There were no significant interaction effects between CORRECTION × ENVIRONMENT, $F_{1,15} = .983, p = .36$ or CURSOR × ENVIRONMENT, $F_{1,15} = 3.79, p = .070$. However, there was a significant interaction between CORRECTION × CURSOR, $F_{1,15} = 4.592, p = .048$, but not between CORRECTION × CURSOR × ENVIRONMENT, $F_{1,15} = 2.03, p = .175$. In summary, using the correction models significantly increases participants' pointing accuracy in the real and in the virtual world. However, the accuracy depends on using a cursor, see Figure 6 and Table 3.

In the following we will estimate target sizes to fit at least 90% of the mid-air pointing actions for all conditions independently. For simplicity we only fit a squared target shape. For *No-Cursor* in RW the sides of the target need to be 17.6 *cm* wide, in VR 18.8 *cm* and with Cursor for RW and VR respectively 4.1 *cm* and 4.5 *cm*. With correction the size for the four squared targets could be respectively 6.9 %, 11.6 %, 6.5 %, and 8.9 % smaller and still fit 90 % of the pointing actions. The estimated target sizes are optimal for a target in 2 *m* distance from the human.

Task completion time (TCT)

We found no significant effects of CORRECTION, $F_{1,15} = .158, p = .697$, or ENVIRONMENT, $F_{1,15} = .004, p = .956$ on the TCT. However, there was a significant effect of CURSOR, $F_{1,15} = 7.834, p = .013$ on TCT. Furthermore, we found significant interaction effects between CURSOR × ENVIRONMENT, $F_{1,15} = 15.61, p < .001$. No interaction effects were found between CORRECTION × CURSOR, $F_{1,15} = .067, p = .799$, between CORRECTION × ENVIRONMENT, $F_{1,15} = 1.291, p = .274$, or CORRECTION × CURSOR × ENVIRONMENT, $F_{1,15} = 1.163, p = .298$. Since the participants received no feedback about their accuracy when using the correction models, the correction model did not affect the TCT in the *real* as well as in the *virtual* environment. However, presenting a cursor increased the time for pointing since the participants used more time to adjust, see Table 2.

Discussion

In our second study, we investigated the effect of the developed models in a real-time setup. As we validated the models for all ray casting techniques only using LOOCV, our evaluation study ensured the external validity of the presented models by inviting 16 new participants. We investigated participants' performance with and without correction models (CORRECTION) as well as the effect of displaying a cursor as pointing indicator on our model (CURSOR). The effect of model and cursor were tested for both real and virtual environments (ENVIRONMENT). As also found in the first study, we found statistically significant differences between *RealWorld*

		No Cursor		*With Cursor*	
	Correction	**M**	**SD**	**M**	**SD**
RealWorld	False	1.48	.43	1.83	.43
RealWorld	True	1.48	.43	1.89	.73
VirtualWorld	False	1.64	.61	1.76	.56
VirtualWorld	True	1.64	.61	1.67	.45

Table 2. Overall TCT to select a the target. TCTs are reported in seconds.

		No Cursor		*With Cursor*	
	Correction	**M**	**SD**	**M**	**SD**
RealWorld	False	7.08	3.26	1.14	.89
RealWorld	True	5.92	3.29	1.13	.96
VirtualWorld	False	6.37	3.42	1.30	.85
VirtualWorld	True	5.76	3.26	1.20	.76

Table 3. Remaining offset interact and target. Distances are reported in cm.

and *VirtualReality*. This supports our choice of building independent models for *RealWorld* and *VirtualReality*, as we found no significant effect of raw TLX over time. Thus, we again assume that the effect of participants' fatigue or workload was negligible.

Our analysis revealed that the offset between the eye-finger ray cast and the target can be significantly decreased in real and virtual environments when using the proposed models. While the models overall improvement without a cursor was 13.1%, the improvement for *VirtualReality* was 9.5% and for *RealWorld* 16.3%. However, the accuracy depends on whether a cursor was displayed or not. With a cursor, the average improvement was 4.5%. The interaction effect of CORRECTION and CURSOR on the accuracy can be explained by a realignment of the user's arm while presenting visual feedback (the cursor) and applying the correcting models. The increased precision is marginally compensated by the user while moving the arm to the target. This is the case in both environments, which is supported by the lacking significant effect of the three-way interaction between CORRECTION, CURSOR, and ENVIRONMENT.

While the accuracy clearly increased when using a cursor which is in line with Cockburn et al. [14], analysis of the TCT revealed that the cursor also increased the time to select a target. However, CORRECTION and ENVIRONMENT did not significantly affect the TCT. Furthermore, the interaction effect of CURSOR and ENVIRONMENT on the TCT was significant. Having a cursor in the real world is potentially less relevant than having a cursor in VR. We assume that this is caused by novelty effects and the users' higher attention to the task while being in VR. The second study shows that the developed models have a positive effect on the mid-air pointing accuracy without a negative effect on the time to select a target. While displaying a cursor also had a positive effect on pointing accuracy, the it also increases the TCT. We, therefore, present the following design considerations for mid-air pointing in both real and virtual environments:

1. Always apply the model to correct systematic mid-air pointing error.
2. For high precise mid-air selection, a cursor should additionally be displayed.
3. For fast mid-air selections, a cursor should not be displayed.

CONCLUSION

In this paper, we built mid-air pointing offset compensation models for real and virtual environments based on pointing gestures of 20 participants. We built models for four different ray casting techniques and used cross-validation (CV) to show that we achieve the smallest remaining offset when using eye-finger ray cast (EFRC). In a second study, we further investigated EFRC in a selection task. We confirm findings of previous work that using a cursor improves mid-air pointing precision. We show that the accuracy of mid-air pointing without a cursor can be improved through correction models for both real and virtual environments by 13.1%. Further, we show that using a cursor a correction model can reduces the remaining pointing error by 4.5%.

As the pointing accuracy may be affected by the HMD we envision as next step a study using HMDs with a variety of FoVs to understand the impact of a limited FoV. In the presented paper we investigated real-world (RW) and virtual reality (VR) which are representing the edges of the Milgram continuum [44], in the next steps, we will also investigate pointing in augmented reality (AR) and mixed reality.

FUTURE WORK

In comparison to Mayer et al. [40] we used a marker set which allowed us to online track the limbs of the participant. We expect that this also contributes towards a more stable and precise tracking. In the future the potential influence of the marker placement should be investigated to determine a universal marker placement. This would contribute towards models which could be applied by everyone who follows the marker placement conventions. This is especially important when future technologies are used for tracking the user without attaching markers but retaining the same precision. On the other hand, this would be also important if the model is applied to already existing less precise tracking technologies like the Microsoft Kinect skeleton tracking.

In both studies the target had no actual size. This was done to build a precise model where there was no room left for the participant to interpret the actual target position. We estimate that the target size can on average be 8.5% smaller when applying our new correction models. Future work should investigate how a target size influences the models' performance.

Incorporating the findings by Plaumann et al. [53] could result in more accurate models and improve pointing accuracy. However, today we cannot determine eye and ocular dominance of a user by just observing the user's behavior. Hence, incorporating eye and ocular dominance would result in user depended models and limit the use cases, e.g. these user-dependent models are not useful for public display scenarios.

ACKNOWLEDGEMENTS

This work was financially supported by the German Research Foundation (DFG) within Cluster of Excellence in Simulation Technology (EXC 310/2) at the University of Stuttgart and through project C04 of SFB/Transregio 161.

REFERENCES

1. Deepak Akkil and Poika Isokoski. 2016. Accuracy of Interpreting Pointing Gestures in Egocentric View. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 262–273. DOI: http://dx.doi.org/10.1145/2971648.2971687

2. Ferran Argelaguet and Carlos Andujar. 2013. A survey of 3D object selection techniques for virtual environments. *Computers & Graphics* 37, 3 (2013), 121 – 136. DOI: http://dx.doi.org/10.1016/j.cag.2012.12.003

3. Ferran Argelaguet, Carlos Andujar, and Ramon Trueba. 2008. Overcoming Eye-hand Visibility Mismatch in 3D Pointing Selection. In *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology*

(VRST '08). ACM, New York, NY, USA, 43–46. DOI: http://dx.doi.org/10.1145/1450579.1450588

4. Ferran Argelaguet, Ludovic Hoyet, Michaël Trico, and Anatole Lécuyer. 2016. The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *Proceedings of the 2016 IEEE Virtual Reality Conference (IEEE/VR '16)*. Greenville, SC, USA, 3–10. DOI:http://dx.doi.org/10.1109/VR.2016.7504682
5. Tiziana Aureli, Maria Spinelli, Mirco Fasolo, Maria Concetta Garito, Paola Perucchini, and Laura D'Odorico. 2017. The Pointing-Vocal Coupling Progression in the First Half of the Second Year of Life. *Infancy* (2017), 18. DOI: http://dx.doi.org/10.1111/infa.12181
6. John V. Basmajian and Carlo J. De Luca. 1985. *Muscles alive: their functions revealed by electromyography*. Williams & Wilkins, Baltimore, London, Sydney. http://books.google.de/books?id=H9pqAAAAMAAJ
7. Scott Bateman, Regan L. Mandryk, Carl Gutwin, and Robert Xiao. 2013. Analysis and comparison of target assistance techniques for relative ray-cast pointing. *International Journal of Human-Computer Studies* 71, 5 (2013), 511–532. DOI: http://dx.doi.org/10.1016/j.ijhcs.2012.12.006
8. Elizabeth Bates. 1979. *The Emergence of Symbols: Cognition and Communication in Infancy*. Academic Press. 404 pages. https://books.google.de/books?id=_45-AAAAMAAJ
9. Hrvoje Benko, Andrew D. Wilson, and Federico Zannier. 2014. Dyadic Projected Spatial Augmented Reality. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 645–655. DOI: http://dx.doi.org/10.1145/2642918.2647402
10. Richard A. Bolt. 1980. Put-that-there: Voice and gesture at the graphics interface. In *Proceedings of the 7th annual conference on Computer graphics and interactive techniques (SIGGRAPH '80)*. ACM, New York, NY, USA, 262–270. DOI: http://dx.doi.org/10.1145/800250.807503
11. Jerome S. Bruner. 1975. The ontogenesis of speech acts. *Journal of Child Language* 2, 1 (1975), 1–19. DOI: http://dx.doi.org/10.1017/S0305000900000866
12. Luigia Camaioni, Paola Perucchini, Francesca Bellagamba, and Cristina Colonnesi. 2004. The Role of Declarative Pointing in Developing a Theory of Mind. *Infancy* 5, 3 (2004), 291–308. DOI: http://dx.doi.org/10.1207/s15327078in0503_3
13. C. N. Christakos and S. Lal. 1980. Lumped and population stochastic models of skeletal muscle: Implications and predictions. *Biological Cybernetics* 36, 2 (01 Feb 1980), 73–85. DOI: http://dx.doi.org/10.1007/BF00361076
14. Andy Cockburn, Philip Quinn, Carl Gutwin, Gonzalo Ramos, and Julian Looser. 2011. Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback. *International Journal of Human-Computer Studies* 69, 6 (2011), 401 – 414. DOI: http://dx.doi.org/10.1016/j.ijhcs.2011.02.005
15. Andrea Corradini and Philip R. Cohen. 2002. Multimodal speech-gesture interface for handfree painting on a virtual paper using partial recurrent neural networks as gesture recognizer. (2002). DOI: http://dx.doi.org/10.1109/IJCNN.2002.1007499
16. Giambattista Della Porta. 1593. *De refractione Optices Parte: Libri Novem...* Ex officina Horatii Salviani, apud Jo. Jacobum Carlinum, & Antonium Pacem.
17. H. Henrik Ehrsson, Charles Spence, and Richard E. Passingham. 2004. That's My Hand! Activity in Premotor Cortex Reflects Feeling of Ownership of a Limb. *Science* 305, 5685 (2004), 875–877. DOI: http://dx.doi.org/10.1126/science.1097011
18. Rodger J. Elble and William C. Koller. 1990. *Tremor*. Johns Hopkins University Press, Baltimore. https://google.com/books?id=-LdrAAAAMAAJ
19. J. M. Foley and Richard Held. 1972. Visually directed pointing as a function of target distance, direction, and available cues. *Perception & Psychophysics* 12, 3 (01 May 1972), 263–268. DOI: http://dx.doi.org/10.3758/BF03207201
20. Edward G. Freedman and David L. Sparks. 2000. Coordination of the eyes and head: movement kinematics. *Experimental Brain Research* 131, 1 (01 Mar 2000), 22–32. DOI:http://dx.doi.org/10.1007/s002219900296
21. Michael Steven Anthony Graziano. 1999. Where is my arm? The relative role of vision and proprioception in the neuronal representation of limb position. *Proceedings of the National Academy of Sciences* 96, 18 (1999), 10418–10421. DOI: http://dx.doi.org/10.1073/pnas.96.18.10418
22. Michael S. A. Graziano, Dylan F. Cooke, and Charlotte S. R. Taylor. 2000. Coding the Location of the Arm by Sight. *Science* 290 (2000), 1782–1786. http://www.jstor.org/stable/3081775
23. Jan Gugenheimer, Pascal Knierim, Christian Winkler, Julian Seifert, and Enrico Rukzio. 2015. UbiBeam: Exploring the Interaction Space for Home Deployed Projector-Camera Systems. In *IFIP TC 13 International Conference on Human-Computer Interaction (INTERACT 2015)*. Springer International Publishing, Cham, 350–366. DOI:http://dx.doi.org/10.1007/978-3-319-22698-9_23
24. Sandra G. Hart. 2006. NASA-Task Load Index (NASA-TLX); 20 years later. In *Proceedings of the Human Factors and Ergonomic Society annual meeting*, Vol. 50. SAGE Publications, SAGE Publications, Los Angeles, CA, USA, 904–908. DOI: http://dx.doi.org/10.1177/154193120605000909
25. John B. Haviland. 2003. *How to point in Zinacantán*. Psychology Press, Mahwah, New Jersey, USA, 139–169.

26. Hao Jiang, Eyal Ofek, Neema Moraveji, and Yuanchun Shi. 2006. Direct Pointer: Direct Manipulation for Large-display Interaction Using Handheld Cameras. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. ACM, New York, NY, USA, 1107–1110. DOI: http://dx.doi.org/10.1145/1124772.1124937

27. Ricardo Jota, Miguel A. Nacenta, Joaquim A. Jorge, Sheelagh Carpendale, and Saul Greenberg. 2010. A Comparison of Ray Pointing Techniques for Very Large Displays. In *Proceedings of Graphics Interface 2010 (GI '10)*. Canadian Information Processing Society, Toronto, Ont., Canada, Canada, 269–276. http://dl.acm.org/citation.cfm?id=1839214.1839261

28. Adam Kendon. 2004. *Gesture: Visible Action as Utterance*. Cambridge University Press, Cambridge. IX, 400 S. pages. https://books.google.de/books?id=WY2bBQAAQBAJ

29. Aarlenne Z. Khan and J. Douglas Crawford. 2003. Coordinating one hand with two eyes: optimizing for field of view in a pointing task. *Vision Research* 43, 4 (2003), 409 – 417. DOI: http://dx.doi.org/10.1016/S0042-6989(02)00569-2

30. Pascal Knierim, Valentin Schwind, Anna Feit, Florian Nieuwenhuizen, and Niels Henze. 2018. Physical Keyboards in Virtual Reality: Analysis of Typing Performance and Effects of Avatar Hands. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA. DOI:http://dx.doi.org/10.1145/3173574.3173919

31. 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 (Oct. 2010), 603–615. DOI:http://dx.doi.org/10.1016/j.ijhcs.2010.05.001

32. Alfred Kranstedt, Andy Lücking, Thies Pfeiffer, Hannes Rieser, and Marc Staudacher. 2006. Measuring and Reconstructing Pointing in Visual Contexts. In *Proceedings of the 10th Workshop on the Semantics and Pragmatics of Dialogue (brandial'06)*. Universitätsverlag Potsdam, Potsdam, 82–89. http://opus.kobv.de/ubp/volltexte/2006/1048/pdf/brandial06_proceedings.pdf

33. Lorraine Lin and Sophie Jörg. 2016. Need a hand? How Appearance Affects the Virtual Hand Illusion. In *Proceedings of the ACM Symposium on Applied Perception (SAP '16)*. ACM Press, New York, New York, USA, 69–76. DOI: http://dx.doi.org/10.1145/2931002.2931006

34. Lars Lischke, Pascal Knierim, and Hermann Klinke. 2015. Mid-Air Gestures for Window Management on Large Displays. In *Mensch und Computer 2015 - Tagungsband*. De Gruyter Oldenbourg, Berlin, 439–442. DOI:http://dx.doi.org/10.1515/9783110443929-072

35. Lars Lischke, Valentin Schwind, Kai Friedrich, Albrecht Schmidt, and Niels Henze. 2016. MAGIC-Pointing on Large High-Resolution Displays. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 1706–1712. DOI: http://dx.doi.org/10.1145/2851581.2892479

36. I. Scott MacKenzie and Shaidah Jusoh. 2001. An Evaluation of Two Input Devices for Remote Pointing. In *Engineering for Human-Computer Interaction: 8th IFIP International Conference (EHCI'01)*. Springer Berlin Heidelberg, Berlin, Heidelberg, 235–250. DOI: http://dx.doi.org/10.1007/3-540-45348-2_21

37. Ville Mäkelä, Tomi Heimonen, Matti Luhtala, and Markku Turunen. 2014. Information Wall: Evaluation of a Gesture-controlled Public Display. In *Proceedings of the 13th International Conference on Mobile and Ubiquitous Multimedia (MUM '14)*. ACM, New York, NY, USA, 228–231. DOI: http://dx.doi.org/10.1145/2677972.2677998

38. Anders Markussen, Sebastian Boring, Mikkel R. Jakobsen, and Kasper Hornbæk. 2016. Off-Limits: Interacting Beyond the Boundaries of Large Displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5862–5873. DOI: http://dx.doi.org/10.1145/2858036.2858083

39. Anders Markussen, Mikkel Rønne Jakobsen, and Kasper Hornbæk. 2014. Vulture: A Mid-air Word-gesture Keyboard. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1073–1082. DOI: http://dx.doi.org/10.1145/2556288.2556964

40. Sven Mayer, Katrin Wolf, Stefan Schneegass, and Niels Henze. 2015. Modeling Distant Pointing for Compensating Systematic Displacements. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 4165–4168. DOI: http://dx.doi.org/10.1145/2702123.2702332

41. Victoria McArthur, Steven J. Castellucci, and I. Scott MacKenzie. 2009. An Empirical Comparison of "Wiimote" Gun Attachments for Pointing Tasks. In *Proceedings of the 1st ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS '09)*. ACM, New York, NY, USA, 203–208. DOI: http://dx.doi.org/10.1145/1570433.1570471

42. Samuel Navas Medrano, Max Pfeiffer, and Christian Kray. 2017. Enabling Remote Deictic Communication with Mobile Devices: An Elicitation Study. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17)*. ACM, New York, NY, USA, Article 19, 13 pages. DOI: http://dx.doi.org/10.1145/3098279.3098544

43. Walter R. Miles. 1930. Ocular Dominance in Human Adults. *The journal of general psychology* 3, 3 (1930), 412–430.

44. Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems* 77, 12 (1994), 1321–1329. http://ci.nii.ac.jp/naid/110003209335/en/

45. Mark R. Mine. 1995. *Virtual Environment Interaction Techniques*. Technical Report. University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. http://www.cs.unc.edu/techreports/95-018.pdf

46. Mark R. Mine, Frederick P. Brooks Jr., and Carlo H. Sequin. 1997. Moving Objects in Space: Exploiting Proprioception in Virtual-environment Interaction. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '97)*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 19–26. DOI: http://dx.doi.org/10.1145/258734.258747

47. Steven Morrison and J. Keogh. 2001. Changes in the dynamics of tremor during goal-directed pointing. *Human Movement Science* 20, 4 (2001), 675 – 693. DOI: http://dx.doi.org/10.1016/S0167-9457(01)00072-0

48. Mathieu Nancel, Emmanuel Pietriga, Olivier Chapuis, and Michel Beaudouin-Lafon. 2015. Mid-Air Pointing on Ultra-Walls. *ACM Trans. Comput.-Hum. Interact.* 22, 5, Article 21 (Aug. 2015), 62 pages. DOI: http://dx.doi.org/10.1145/2766448

49. Kai Nickel and Rainer Stiefelhagen. 2003. Pointing Gesture Recognition Based on 3D-tracking of Face, Hands and Head Orientation. In *Proceedings of the 5th International Conference on Multimodal Interfaces (ICMI '03)*. ACM, New York, NY, USA, 140–146. DOI: http://dx.doi.org/10.1145/958432.958460

50. Dan R. Olsen, Jr. and Travis Nielsen. 2001. Laser Pointer Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '01)*. ACM, New York, NY, USA, 17–22. DOI: http://dx.doi.org/10.1145/365024.365030

51. Jeni Paay, Dimitrios Raptis, Jesper Kjeldskov, Mikael B. Skov, Eric V. Ruder, and Bjarke M. Lauridsen. 2017. Investigating Cross-Device Interaction Between a Handheld Device and a Large Display. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 6608–6619. DOI: http://dx.doi.org/10.1145/3025453.3025724

52. Jeffrey S. Pierce, Andrew S. Forsberg, Matthew J. Conway, Seung Hong, Robert C. Zeleznik, and Mark R. Mine. 1997. Image Plane Interaction Techniques in 3D Immersive Environments. In *Proceedings of the 1997 Symposium on Interactive 3D Graphics (I3D '97)*. ACM, New York, NY, USA, 39–ff. DOI: http://dx.doi.org/10.1145/253284.253303

53. Katrin Plaumann, Matthias Weing, Christian Winkler, Michael Müller, and Enrico Rukzio. 2017. Towards accurate cursorless pointing: the effects of ocular dominance and handedness. *Personal and Ubiquitous Computing* (07 Dec 2017). DOI: http://dx.doi.org/10.1007/s00779-017-1100-7

54. Cameron N. Riviere, R. Scott Rader, and Pradeep K. Khosla. 1997. Characteristics of hand motion of eye surgeons. In *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE/EMBS '97)*, Vol. 4. IEEE, Chicago, IL, USA, USA, 1690–1693. DOI: http://dx.doi.org/10.1109/IEMBS.1997.757046

55. Valentin Schwind, Pascal Knierim, Cagri Tasci, Patrick Franczak, Nico Haas, and Niels Henze. 2017. "These Are Not My Hands!": Effect of Gender on the Perception of Avatar Hands in Virtual Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1577–1582. DOI: http://dx.doi.org/10.1145/3025453.3025602

56. Mel Slater, Daniel Perez-marcos, and H. Henrik Ehrsson. 2009. Inducing illusory ownership of a virtual body. *Frontiers in neuroscience* 3, 2 (Sep 2009), 214–220. DOI: http://dx.doi.org/10.3389/neuro.01.029.2009

57. 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. DOI: http://dx.doi.org/10.1016/j.ijhcs.2004.09.001

58. John S. Stahl. 1999. Amplitude of human head movements associated with horizontal saccades. *Experimental Brain Research* 126, 1 (01 Apr 1999), 41–54. DOI:http://dx.doi.org/10.1007/s002210050715

59. Daniel Vogel and Ravin Balakrishnan. 2004. Interactive Public Ambient Displays: Transitioning from Implicit to Explicit, Public to Personal, Interaction with Multiple Users. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology (UIST '04)*. ACM, New York, NY, USA, 137–146. DOI: http://dx.doi.org/10.1145/1029632.1029656

60. Daniel Vogel and Ravin Balakrishnan. 2005. Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology (UIST '05)*. ACM, New York, NY, USA, 33–42. DOI:http://dx.doi.org/10.1145/1095034.1095041

61. Christian Winkler, Ken Pfeuffer, and Enrico Rukzio. 2012. Investigating Mid-air Pointing Interaction for Projector Phones. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces (ITS '12)*. ACM, New York, NY, USA, 85–94. DOI:http://dx.doi.org/10.1145/2396636.2396650

62. Nelson Wong and Carl Gutwin. 2014. Support for Deictic Pointing in CVEs: Still Fragmented After All These Years'. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. ACM, New York, NY, USA, 1377–1387. DOI: http://dx.doi.org/10.1145/2531602.2531691

63. Shumin Zhai. 1998. User Performance in Relation to 3D Input Device Design. *ACM SIGGRAPH Computer Graphics* 32, 4 (Nov. 1998), 50–54. DOI: http://dx.doi.org/10.1145/307710.307728