Up to the Fingertip: The Effect of Avatars on Mid-Air Pointing Accuracy in Virtual Reality

Valentin Schwind[1], **Sven Mayer**[1], **Alexandre Comeau-Vermeersch**[2],
Robin Schweigert[1], **Niels Henze**[1,3]
[1]University of Stuttgart, Stuttgart, Germany, {valentin.schwind, sven.mayer,
robin.schweigert}@vis.uni-stuttgart.de
[2]Université de Sherbrooke, Sherbrooke, Canada,
Alexandre.Comeau-Vermeersch@USherbrooke.ca
[3]University of Regensburg, Regensburg, Germany, niels.henze@ur.de

ABSTRACT

Avatars in virtual reality (VR) increase the immersion and provide an interface between the user's physical body and the virtual world. Thus, avatars enable referential gestures, which are essential for targeting, selection, locomotion, and collaboration in VR. However, players of immersive games can have another virtual appearance deviating from human-likeness and previous work suggests that avatars can have an effect on the accuracy of referential gestures in VR. One of the most important referential gestures is mid-air pointing. It has been shown that mid-air pointing is affected by systematic errors, which can be compensated using different methods. Thus, it is unknown if the avatar must be considered in corrections of the systematic error. In this paper, we investigate the effect of the avatar on pointing accuracy. We show that the systematic error in pointing is significantly affected by the virtual appearance but does not correlate with the degree to which the appearance deviates from the perceived human-likeness. Moreover, we confirm that people only rely on their fingertip and not on their forearm or index finger orientation. We present compensation models and contribute with design implications to increase the accuracy of pointing in VR.

CCS Concepts

•**Human-centered computing** → **Virtual reality; Pointing;**
•**Computing methodologies** → *Virtual reality;*

Author Keywords

Virtual reality; avatars; pointing; ray-casting; accuracy

INTRODUCTION

Virtual reality (VR) provides immersive experiences and even allows users to take on and control virtual bodies. The virtual body ownership illusion occurs when the avatar is rendered according to the user's pose and movement in the real world. Such avatars are fundamental design elements of games and other immersive applications in VR. Therefore, it is important to understand how virtual avatars and the user's own body can be used as gaming controllers and how the virtual avatar is integrated into the own body scheme.

Figure 1. User with the marker set of rigid bodies for finger, forearm, upper arm, head, left, right shoulder, and head (HMD).

Previous work found that the appearance of the avatar impacts the illusion of virtual limb ownership. The human-likeness of an avatar increases the degree of limb ownership and the user's sense of agency and behavior in VR [3]. Thus, the similarity of the avatar with the human appearance can increase the illusion of body ownership. However, the realism of the virtual body can also negatively affects the acceptance of the virtual appearance and causes a disconnection from the virtual body. For example, the wrong gender [40], mismatching limbs [39], or even too realistic avatars [26, 46] produce uncomfortable feelings and altered behavior. Some of these findings were partially brought into connection with a visually induced dysphoria [39, 40] or the uncanny valley phenomenon [26, 46], which describes a sudden reduction in familiarity towards very human- [32] or animal-like [41] characters.

It has been shown that avatars not only influence the subjective experience or the feeling of virtual body ownership in VR. The avatar can also influence the task performance [22] and the user's movements [3]. Furthermore, researchers found that the

CHI PLAY '18, October 28–31, 2018, Melbourne, VIC, Australia

DOI: https://doi.org/10.1145/3242671.3242675

rendering of the own avatar can affect the workload and typing performance in VR [23]. However, task performance or behavioral changes have been observed without considering the accuracy of the performed interaction or gesture. To provide interaction techniques for VR and games without additional hand controllers, it is not only essential to render the desired avatar appearance, but also to learn how the avatar influences the accuracy when performing gestures.

Referential gestures require high accuracy and are essential for interacting in VR. They are used for targeting, selection, locomotion, and collaboration in VR. The mid-air pointing gesture is a universal activity and this "primeval referential device" [17] is developed in early childhood, and plays a fundamental role in developing a theory of mind [9]. Furthermore, the gesture depends on the ability of self-location and the coordination of head, arms, hands, and fingers. Despite intensive efforts to point precisely at targets, humans make systematic errors while pointing [13, 27, 28]. As systematic errors can be compensated, researchers are interested in models to resolve inaccuracies when interpreting the users' input.

Using context and deduction, humans are well skilled in overcoming potential inaccuracies or ambiguities while mid-air pointing at a specific target. However, when interacting with virtual systems, methods are required to determine the precise pointing direction. Current systems use ray casting methods based on finger, hand, and body pose tracking [2, 10, 35]. Systematic error compensation models can improve the direction derived by these ray casting techniques [28, 27]. Using compensating models is the state-of-the-art approach of systems detecting pointing directions as referential inputs, which are not only used in immersive VR games but also for interactions with large high-resolution displays [20, 47], in media rooms [7], and in smart home environments [19]. Immersive applications such as games can benefit from offset correction models as such models can increase the accuracy of target selection [48], locomotion in virtual environments [44], and deictic pointing gestures in multi-user environments [37]. While systematic errors can easily be corrected with corresponding models and ray casting methods, the virtual avatar in games can vary or even disappear and, thus, may distort and falsify corresponding corrections.

As avatars can have an effect on the input performance in VR [23] and point [27] differently in the real world compared to VR, we hypothesize that the appearance of avatars will not equally be integrated into the own body scheme while performing referential gestures using different avatars. Similar effects of the virtual hand appearance are also evident in visual-haptic integration [42]. Current VR applications can provide continuous feedback to support pointing gestures. However, immersive games such as *Skyrim VR* [6], *Arizona Sunshine* [45], and *Fallout 4 VR* [5] do not provide continuous feedback. Additional visual feedback can break the VR experience in realistic games and increase the cognitive load [48]. Thus, we propose to use correcting models without additional feedback. However, general models reducing the pointing error run into danger to be no longer valid in VR as the avatar can have a complete different appearance.

Based on previous work in the research domains of HCI and virtual limb ownership, we hypothesize that people represented by less human-like avatars will not ignore the visual cue given by a virtual avatar with a deviating appearance from the own one and, thus, increase the accuracy of a mid-air pointing gesture. To test this hypothesis, we conducted a pointing study in VR using different whole-body avatars. Furthermore, we hypothesize that people mainly rely on the direction of their fingertip ray cast. Our contribution is three-fold:

- We show that avatars have a significant effect on pointing accuracy. For example, robot and abstract avatars can increase pointing accuracy compared to human-like avatars.
- We found that people mainly rely on the position of their fingertip when pointing and not on the direction of their forearm or index finger.
- We provide polynomial models for the herein presented avatars to compensate mid-air pointing displacements and significantly reduce the systematic pointing error.

RELATED WORK

This paper is related to previous work investigating the effects of virtual limb- and body-ownership on perception and behavior. Our work is based on investigations of the pointing gesture in Psychology, Physiology, and human-computer interaction (HCI). Researchers compare the accuracy of ray casting methods, and compensate the error with polynomial displacement models. Our work is related to absolute pointing without using additional devices [34] or visual feedback [47] and only partially refers to relative pointing techniques [7, 8].

Virtual Limb- and Body-Ownership

As mentioned prior research found that the brain is able to accept virtual limbs [12] and bodies [43] as parts of the own body. Rendering a virtual body solves a fundamental problem of self-location as the brain encodes limb position primarily using vision [14, 15]. However, it has also been shown that the appearance of an avatar affects the illusion of body ownership. For example, Lin and Jörg [25] found that human-like hand models increased the illusion of body ownership. Vinayagamoorthy et al. [46] and Lugrin et al. [26] found higher levels of virtual limb ownership in VR when using fewer realistic VR game characters. The authors of both papers assume that presence is affected by the uncanny valley phenomenon by Mori [32], who hypothesized that imperfections of very human-like characters cause uncomfortable sensations. Schwind et al. [40] found gender-related difference and, for example, recommend to avoiding gender swapping in VR by using non-realistic or androgyny avatars.

Changing the avatar appearance can also lead to behavioral changes. Argelaguet et al. [3] found that the sense of agency was stronger for less human-like virtual hands, however, the illusion of body ownership increased with human-like virtual hands. Researchers also showed that not only the perception but also typing performance with avatar hands depends on their appearance [23]. In a study with less fingers, Schwind et al. [39] observed that users only used fingers they saw for interaction. Furthermore, the authors found that the feeling of presence in VR depends on structural changes and the realism of avatar hands. Thus, the illusion of virtual ownership is significantly affected by the virtual appearance and can lead to changes in behavior, task performance, and presence in VR. Currently, knowledge is lacking about the accuracy while interacting or pointing with different avatars in VR.

Pointing as Referential Gesture

Research in psychology has shown that children begin to express themselves with pointing gestures in early childhood [17]. The pointing gesture helps humans to learn other's intentions and has an impact on developing a theory of mind [9] as well as on learning verbal declarations [4]. In his book, Kendon [21] differentiates pointing gestures using the index finger, open hand, or the 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 being meant [21]. Pointing requires fine levels of dexterity and control over intrinsic oscillations of the own body (tremor). Mid-air pointing is also affected by the tremor of the human muscle system. Morrison and Keogh [33] conducted a frequency analysis for pointing with the hand and index finger and found two dominant frequency peaks at $2-4$ and $8-12$ Hz. They also found that oscillations increased when participants attempted to reduce the tremor by exerting greater control over the hand.

Previous work proposes multiple ray casting approaches to compute the pointing direction. Argelaguet et al. [2] distinguish between two families of ray casting methods: *hand-* and *eye-rooted* techniques. Using hand-rooted methods the ray from the hand differs from the eye/head position and can intersect with objects even when there is no visible occlusion of the target for the eye (or vice versa) [30, 31]. Corradini and Cohen [10] identified the index finger ray cast (IFRC) as the most common hand-rooted method. For eye-rooted ray casting approaches, there are two methods: the ray cast direction given by the eyes' orientation (gaze ray casts) [35] or the ray between eye and index finger tip position. Kranstedt et al. [24] suggest using the ray between the eyes ('Cyclops Eye') and the index finger tip. According to Mayer et al. [28] we refer to this approach as eye-finger ray cast (EFRC). Nickel et al. [35] proposed and investigated an *elbow-rooted* ray casting method. We refer to this method as forearm ray cast (FARC). Jota et al. [20] recommend reducing the parallax and considering the kind of task as well as the system for ray cast selection.

Foley et al. [13] found a distance dependent trend while pointing with the index finger. The pointing gesture overreaches targets to the opposite side of the dominant hand and sighting eye. This was confirmed by Mayer et al. [28]. In their work, the authors describe the systematic error in absolute pointing and present a polynomial displacement model for compensation. Similarly, Akkil and Isokoski [1] conducted an experiment to compare different pointing techniques including eye gaze to compensate the error. Their results indicate that overlaying gaze information on an egocentric view increase the accuracy and confidence while pointing. In another study by Mayer et al. [27], which is closely related to our work, the authors showed a difference between the systematic error in the real and virtual world. They used a compensating model and a cursor to improve the users' accuracy in pointing; however, their study was only validated using a human-like avatar. It is conceivable, that further avatars exist that produce a different error and, thus, require different compensating models.

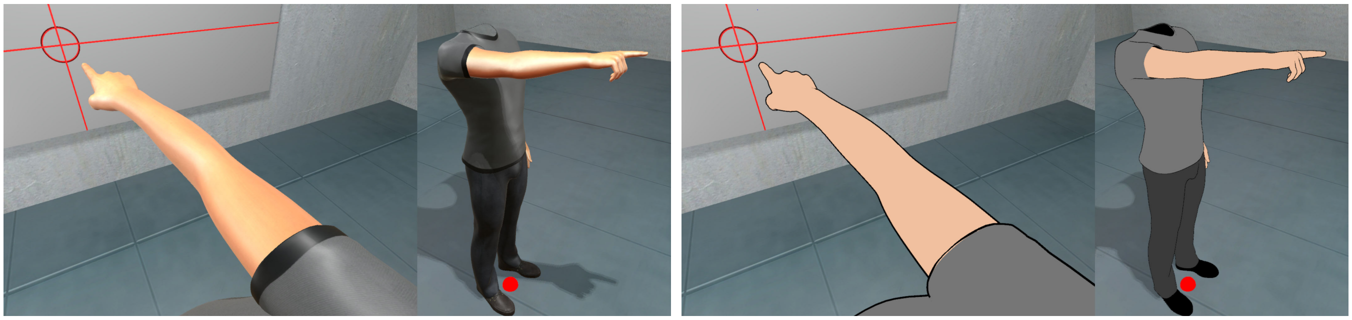
Summary

As shown by previous work, a virtual avatar increases the sense of virtual limb- and body-ownership [43] and improves self-localization [14, 15]. However, different virtual avatars are not perceived equally [26, 40, 46] and cause perceptual as well as behavioral changes [3, 23, 25]. Mayer et al. [27] showed a systematic difference of the systematic error in pointing between the real and the virtual environments, but they did not investigate the effect of the avatar representation itself.

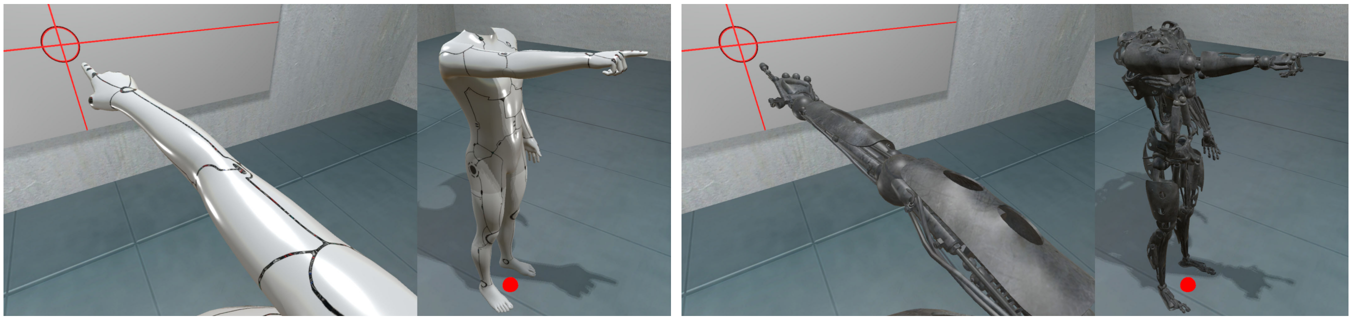
STUDY

Mir-air pointing has the potential to enrich the gaming experience in VR. Object selection, locomotion, and communicating with other players or even AI-avatars, are examples of the large number of use cases. To enable game designers to use mid-air pointing in-game it is important to understand if and how the avatar affects pointing performance. Therefore, we compare different AVATARS to investigate how self representations affect humans' mid-air pointing accuracy in VR. While Mayer et al. [27] showed a difference in the systematic error when pointing in real and the virtual worlds, they only used a single upper body avatar to represent the participant in VR. The avatar was human-like, however, it has been shown that avatar perception varies with the human-likeness of the character [3, 40, 46]. Thus, we hypothesize that users in VR rely only on informative cues of their own body to perform pointing gestures, which potentially increases their accuracy as they cannot ignore an avatar deviating from their own appearance. Furthermore, when people are used to relying on the visual cues given by their body, the mean error should increase when no avatar is visible. Our methodology for measuring and compensating the systematic error [13] mainly builds upon the work by Mayer et al. [27, 28].

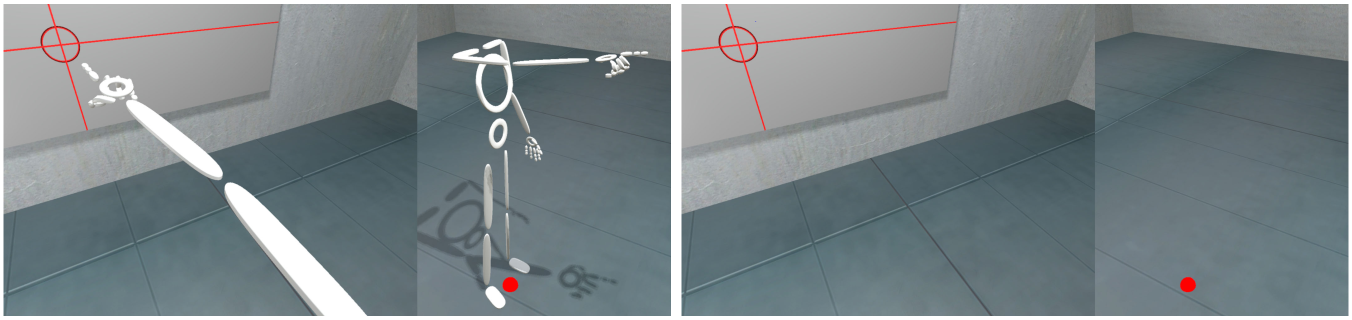
The stimuli selection is based on previous work, investigating the human-likeness or realism of avatars in VR [3, 25, 40]. We used avatars in different rendering styles and varied the degree of human-likeness by changing the texture as well as geometrical morphology of whole body avatars. We used a within-subject design with AVATAR as the only independent variable with 6 levels: *human*, *cartoon*, *robo suit*, *robot*, *abstract* and *invisible*, as shown in Figure 2. The dependent variable is the distance between the intersect of the ray cast on a virtual screen and the targets. We used three different ray casting techniques as purposed by previous work: EFRC, IFRC, and FARC. As proposed by previous work [39, 40], we asked participants to assess human-likeness, attractiveness, and eeriness of each AVATAR. Since the tasks might be tiring for some subjects, we asked for potential workload and fatigue effects using a raw NASA-Task Load Index (raw TLX) between the conditions.



(a) *Human* (b) *Cartoon*



(c) *Robo Suit* (d) *Robot*



(e) *Abstract* (f) *Invisible*

Figure 2. First- VR and third-person view of the six avatars pointing at one of the virtual targets: *Human* (a), *Cartoon* (b), futuristic *Robo Suit* (c), a mechanical *Robot* (d), an *Abstract* (e) and *Invisible* (f) avatar. All avatars used the same skeleton. The avatars in the first row (a-c) used the same human model and skinning. Body parts of the mechanical *Robot* and the *Abstract* avatars used rigid skinning. Red point on the floor indicate the target standing position for the participants.

Apparatus

We used a Windows 10 PC with an NVIDIA GeForce GTX 1080 connected to an OptiTrack system with 14 cameras to determine the position of 18 markers placed over the participant's upper body and right arm. These markers were used to track the exact position of the participant's left and right shoulders, upper arm, forearm, hand and index finger. The position of the body was moved according to the center of both shoulders. Additional body parts were not animated.

OptiTrack offers a commercial implementation and approximation of a full body skeleton, however, it does not support index finger tracking. Thus, we used rigid bodies to track the upper right side of body without any body approximations through the OptiTrack system as suggested in previous work [27]. We used 3D printed custom mounts by Mayer et al. [27] to fit the shape of the arm, hand, finger, and HTC Vive HMD[1]. Index finger markers wrapped around the finger. The upper and forearm marker mounts are wrapping around the arm with hook and loop fastener. The marker setup is shown in Figure 1.

The virtual scene presented a medium sized room, as one would find in a first-person game, for example, equipped with concrete textures. The volume of the virtual room was $6m \times 5.55m \times 3.5m$ (length × width × height). The virtual screen was a wide blank canvas on which the targets were positioned $2m$ away in the front of the participant. The VR scene was rendered using Unity 3D v5.6 and an HTC Vive HMD with a refresh rate of 90 frames per second and a field of view of 110. Positional tracking provided by the HTC Vive was not used. The virtual apparatus was scaled and arranged in real

---

[1] `https://github.com/interactionlab/htc-vive-marker-mount`

world units (meter). To create results which are comparable to previous work we decided to use the same target arrangement as Mayer et al. [27]. Thus, the targets were arranged in a 7×5 (column × row) grid. The targets were presented in VR on a grey virtual screen. The spacing of the target grid was $44.9cm \times 34.cm$.

Stimuli

We used whole body avatars as referential device to avoid distracting participants through cut off or levitating limbs. To avoid further distractions from the animation of other body parts, only their upper body and the right arm were tracked and animated. The stimuli were carefully selected to learn more about potential biases caused by morphology and rendering of human-like avatars. All avatars are based on the Genesis 3 Character by DAZ3D[2]. As suggested by Schwind et al. [40], we removed gender specific cues (*e.g.,* hairs, and glossy nails). For the *human*, *cartoon*, and *robo suit* avatars, we used the realistic human-like model of the generic Genesis 3 character by DAZ3D. To avoid potential distractions by the rest of the body, the *human* and *cartoon* avatar received a thin T-shirt, pants, and shoes.

The *cartoon* model received three toon shader materials for skin (skin color), pants (dark grey), and the T-shirt (grey). The *robo suit* avatar received no clothing because of the futuristic and glossy suit texture. The *robot* (the "Cypher" 3D model and subclass of Genesis 3) and *abstract* avatars are based on the design by Argelaguet et al. [3] and used solid rigid body skinning of the mechanical body parts. No body parts or referential devices were rendered in the *invisible* AVATAR condition. To ensure that obviously affecting cues of the virtual body such as position, size, and length of the limbs do not compromise our measure, we used the same skeleton and the same limb lengths throughout for all avatar geometries.

In line with [27], the avatars were scaled to match the individual subjects' height and size. The animation of *all* models referred exactly to the same body skeleton given by DAZ3D. The *human*, *cartoon*, and *robo suit* avatars had the same skin mesh topology and skinning weights. As participants had to refer to virtual targets using their finger tip, all meshes had the same finger length. The *abstract* and *robot* avatars had a different morphology but the same length of arm and finger as the *human*, *cartoon*, and *robo suit* avatars. The mechanical body parts of the *robot* and the primitives in the *abstract* model were placed according to the orientation of the skeleton bones from DAZ3D. Thus, only the textural (Figures 2a to 2c) or the morphological structure of the rest of the body (Figures 2d to 2f) were changed as referential device.

Measures

We measured the absolute distance of ray cast intersections with the virtual screen. As suggested by related work, we measured the offset between hit and target using EFRC, IFRC, and FARC. Furthermore, we used the raw TLX questionnaire [16] to assess potential fatigue or work load effects. To assess the subjective perception of the avatars we used the questionnaire proposed by Ho and MacDorman [18] to assess the human-likeness, attractiveness, and eeriness of each condition.

Task

For the pointing task, we followed the description by Mayer et al. [28]. The virtual targets were placed in an invisible uniformly distributed 7×5 grid (column × row) for a total of 35 targets in front of the participant. The virtual screen had a size of $269.4cm \times 136.2cm$. The spacing of the target grid was $44.9cm \times 34.cm$. An encircled red cross was presented at each of the 35 positions. The order of appearance of the targets was randomized for each participant. Participants were required to point at each target position twice per condition for a total of 420 targets for the whole experimental session. Furthermore, the participants were equipped with a wireless presenter in their left hand to start the sample recording while they are pointing.

To compensate natural hand tremor, stated as an issue by Olsen and Nielsen [36], the participants had to hold their pointing pose for one second. The participants had to click with the left hand on a button of the remote control when they started holding and the target disappeared after one second. We asked participants to point as they would naturally do in other situations. They had to take their arm down after each recording, and wait from 2 to 3 seconds until they were asked to point at the next target.

Procedure

After welcoming, all participants were introduced with a brief description of the procedure and the purpose of the study. After signing the consent form, they were equipped with tracking markers and the VR HMD. We took 14 measurements to have a real-scale representation of the participant in VR. At the beginning of each condition, the participants walked to the center point of the tracking volume indicated by a red point on the virtual floor with a distance of $2m$ from the screen. AVATARS were sorted in a 6×6 Latin Square design. After each condition, the participants took off the VR headset, sat down, and completed the questionnaires.

Participants

A total of 24 participants (16 male, 8 female) from Central Europe and North America took part in our study. The mean age was 21.25 years ($SD = 1.98$). Findings in prior work indicate that taking eye dominance into account can improve the accuracy in interpreting the directions of the pointing gesture. In line with Plaumann et al. [38], we used the Miles [29] and Porta test [11] to screen participants for eye dominance. Fourteen participants had right-eye dominance, 8 left-eye dominance, and 2 were unclear. Six participants used vision correction during the study. All participants were right-handed. Participants were compensated with credit points for participation in their class.

RESULTS

Participating in the study took $M = 58.12$ minutes ($SD = 9.75$) per participant. Average time for each condition was $M = 5.20$ ($SD = 1.19$) for the *human* avatar, $M = 4.95$ ($SD = 1.31$) for the *cartoon* avatar, $M = 5.14$ ($SD = 1.31$) for the

---

[2] https://www.daz3d.com

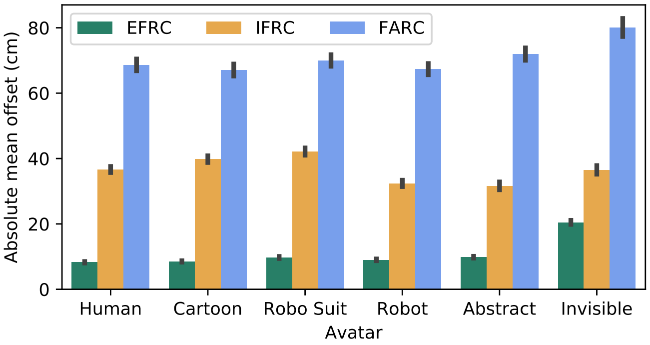


Figure 3. The mean distance between the intersection of the ray cast and the target (in cm) of the six avatars using the three ray casting methods (EFRC, IFRC, FARC). Error bars show 95% confidence intervals (CI_{95}).

abstract avatar, $M = 5.10$ ($SD = 1.16$) for the *robo suit*, $M = 5.10$ ($SD = 1.28$) for the *robot*, and $M = 4.98$ ($SD = 1.02$) for the *Invisible* condition. Each participant performed 420 pointing gestures (70 in each condition). In line with Mayer et al. [27], we report the errors in meters to understand the actual error/influence on any VR scenario. Using a static $2m$ projection canvas we ensure comparability with previous work. Furthermore, we used the instructions by Mayer et al. [27] to determine the ray casts.

Potential Fatigue/Workload Effects

First, we analyzed the raw TLX score to determine if potential workload or fatigue effects had to be considered in the analysis. The mean raw TLX score was $M = 45.48$ ($SD = 24.48$) after the first, $M = 38.61$ ($SD = 25.54$) after the second, $M = 41.66$ ($SD = 27.34$) after the third, $M = 39.58$ ($SD = 24.99$) after the fourth, $M = 36.87$ ($SD = 27.34$) after the fifth, and $M = 38.40$ ($SD = 29.55$) after the last phase. A one-way repeated measures analysis of variance (RM-ANOVA) was conducted. As the analysis did not reveal a significant effect, $F(5,23) = .918, p = .471$, we assume that the effect of participants' fatigue or workload is negligible.

Distance Offset Analysis

We used the ray casting methods EFRC, IFRC, and FARC to measure the distance offset between the intersection of the pointing ray and the actual target on the virtual screen. We removed outliers for each ray casting method, condition, and target individually. Samples more than three standard deviations away from the average of the center were omitted (15 EFRC, 205 IFRC, 2120 FARC). Resulting in a total of 37,750 samples which were considered in the analysis of the ray casting methods. All offset means are listed in Table 1 and depicted in Figure 3. As comparison, Mayer et al. [27] report a $1.05cm$ (9.3%) smaller offset in IFRC and $6.56cm$ (18.4%) smaller offset in IFRC, however, a $4.81cm$ (8.7%) larger offset in FARC.

First, we conducted a repeated measure (RM) one-way multivariate analysis of variance (MANOVA) to determine if all three ray casting methods were altered by the manipulation of the independent variable (Avatars). Using Wilks' lambda, we found a significant effect, $\Lambda = .115, F(3,15) = 4.613$, $p = .013$, $\eta^2 = .885$. Separate univariate analysis of variances (ANOVAs) on the dependent variables revealed significant effects on EFRC, $F(5,115) = 48.345$, $p < .001$), IFRC, $F(5,115) = 5.293$, $p = .002$), however, not on FARC, $F(5,115) = .159$, $p = .159$).

Bonferroni-corrected pairwise comparisons revealed significant differences of the distance error using EFRC between the *invisible* and the other levels of Avatar (all with $p < .001$). No differences were found between the other levels (all with $p = 1.000$). For the IFRC, we found significant differences between the *cartoon* and *abstract* ($p < .001$), the *human* and *abstract* ($p = .031$), *cartoon* and *invisible* ($p = .031$), *cartoon* and *robot* ($p = .021$), *human* and *robot* ($p = .031$), and between *robot* and *robo suit* ($p < .001$). All means are listed in Table 1 and depicted in Figure 3.

Two-dimensional Offset Analysis

The offset analysis only considers the one-dimensional distance between ray intersection and actual target on the virtual screens. However, to understand the direction of displacements it is also important to know if both axes of a vector have to be considered independently from each other and if they are affected by the avatars and the arm or finger movement. Therefore, we conducted a multivariate multiple regression analysis using a mixed effects model with the *horizontal* and *vertical offset* as the two latent dependent variables indicating the direction of the vector displacement. Horizontal, Vertical were used as continuous covariates given by their position on the target grid. Ray Cast and Avatars were used as exploratory variables of the model. By regarding Ray Cast as predictor (not as kind of measurement) we learn how the offset was influenced by arm and finger poses.

The multivariate regression equation with x- and y-offsets as dependent measures was significant, $R^2 = .433, R^2_{Adj.} = .432, SE = .162, F(19,14146) = 568.9, p < .001$. Difference to the mixed model without Avatars as a exploratory factors ($R^2 = .372, R^2_{Adj.} = .371$) was significant (p < .001). The results of the complete model were significant for Avatar, Pillai's Trace = .067, $F(10,14146) = 98.4$, $p < .001$, $\eta^2 = .035$, Ray Cast, Pillai's Trace= 1.010, $F(4,14146) = 7224.3$, $p < .001$, $\eta^2 = .506$, Horizontal, Pillai's Trace= .032, $F(2,14146) = 236.8$, $p < .001$, $\eta^2 = .032$, Vertical, Pillai's Trace= .006, $F(2,14146) = 48.0$, $p < .001$, $\eta^2 = .006$, and Avatar × Ray Cast, Pillai's Trace= .046,

	EFRC		IFRC		FARC	
Avatar	M	SD	M	SD	M	SD
Human	9.15	6.56	37.41	20.77	54.49	21.13
Cartoon	9.31	6.3	40.29	21.74	53.73	21.2
Robo Suit	9.71	6.57	39.28	20.06	56.04	21.62
Robot	9.39	6.42	33.29	21.12	55.14	21.23
Abstract	9.82	5.73	30.14	21.06	56.18	22.21
Invisible	20.19	11.66	33.77	20.24	56.53	20.06
Average	11.26	7.21	35.7	20.83	55.35	21.24
Mayer et al.[27]	8.45	4.54	27.03	18.95	69.66	25.42

Table 1. Absolute means (M) and standard deviations (SD) of the distance between intersection of the ray cast and target (in *cm*).

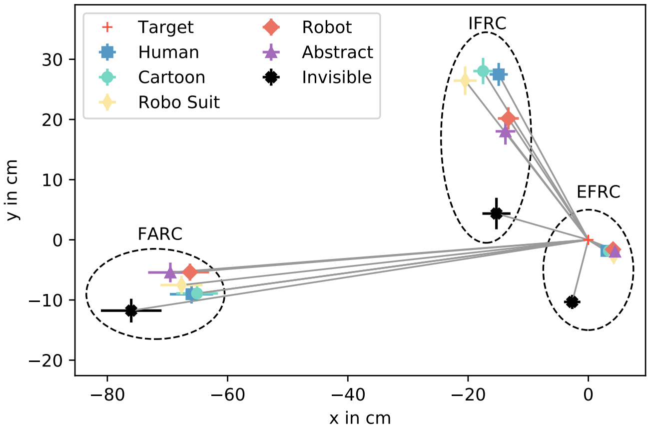


Figure 4. The vector plots for all ray casting methods (EFRC, IFRC, FARC) averaged over all targets. Error bars show 99% confidence intervals (CI_{99}).

$F(20,14146) = 33.8$, $p < .001, \eta^2 = .023$. We learned that both axes of a vector have to be considered horizontal as well as in vertical. To understand how each of the two directions was influenced by our factors, we conducted two separate univariate mixed-model analyses. All scatterplots (not illustrated) of standardized residuals indicated that the data met the assumptions of homogeneity of variance, linearity, and homoscedasticity for all regression analyses.

The univariate regression equation for the x-axis was significant, $R^2 = .635, R^2_{Adj.} = .634, SE = .175, F(19,14146) = 1296, p < .001, d = 1.949$. No auto-correlations d were found ($\rho = -.025$, $p = .004$). For offsets on the x-axis, we found a significant effects for AVATAR, $F(5,14146) = 5.071$, $p < .001$, $\eta^2 = .002$, RAY CAST, $F(2,14146) = 11976.9$, $p < .001$, $\eta^2 = .632$, HORIZONTAL, $F(1,14146) = 465.6$, $p < .001$, $\eta^2 = .032$, VERTICAL, $F(1,14146) = 11.5$, $p < .001$, $\eta^2 < .001$, and AVATAR × RAY CAST, $F(10,14146) = 16.935$, $p < .001$, $\eta^2 = .011$. The univariate regression equation for the y-axis was also significant, $R^2 = .433$, $R^2_{Adj.} = .432$, $SE = .162$, $F(19,14146) = 568.9$, $p < .001$, $d = 1.917$, and no auto-correlations d were found ($\rho = 0.041$, $p < .001$). For the offsets between the interact of the ray cast and the target on the y-axis, we found a significant effect for AVATAR, $F(5,14146) = 183.3$, $p < .001$, $\eta^2 = .063$, RAY CAST, $F(2,14146) = 4624.9$, $p < .001$, $\eta^2 = .394$, HORIZONTAL, $F(1,14146) = 3.7$, $p = .050$, $\eta^2 < .001$, VERTICAL, $F(1,14146) = 67.7$, $p < .001$, $\eta^2 = .038$, and AVATAR × RAY CAST, $F(10,14146) = 57.193$, $p < .001$, $\eta^2 = .038$.

Through multi- and univariate variance analyses using mixed effects models, we learned that the avatar influenced the offsets in our pointing experiment for all ray casting techniques as shown by the interaction effects of the models. The results also show that the variance of x- and y-offsets is independent from each other and do not homogeneously tend towards a single direction. The vector plots of all offsets and ray casting methods are shown in Figures 4 and 6.

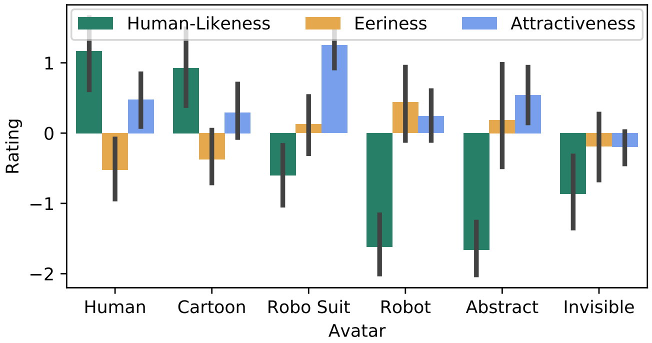


Figure 5. Subjective ratings of the perceived human-likeness, attractiveness, and eeriness for each avatars. Error bars show CI_{95}.

Subjective Perception of the Avatars

We conducted three one-way RM-ANOVA to assess the perceived human-likeness, attractiveness, and eeriness of each avatar. We used the uncanny valley questionnaire by Ho and MacDorman [18] and found significant effects on human-likeness, $F(5,138) = 24.719$, $p < .001$, on attractiveness, $F(5,138) = 6.696$, $p < .001$, however, not on eeriness, $F(5,138) = 1.949$, $p = .09$. Thus, we assume that potential effects of the uncanny valley as found by prior work are negligible.

Pairwise Bonferroni-corrected post-hoc comparisons of the perceived human-likeness measures revealed significant effects between all measures (with $p < .05$), except between *human* and *cartoon*, *abstract* and *robot*, *robo suit* and the *invisible*, *robot* and *invisible* Avatar (with $p > .05$). Pairwise comparisons of the perceived attractiveness showed significant differences between the *cartoon* and *robo suit*, *abstract* and *invisible*, as well as between the *robo suit* and *invisible* Avatar (all other with $p > .05$). Multiple linear regression analyses were conducted to understand if the perceived measures potentially correlates with the quantitative measures. None of the regression equations was significant (all with $p > .05$). All means ratings of the questionnaire are depicted in Figure 5.

Model Development for Compensation and Prediction

Mayer et al. [28] showed that the presented systematic offsets can be correct using two-dimensional polynomial functions with the offset angles α_{pitch} and α_{yaw} as input. They proposed using two functions one to model the offset in pitch direction and one for the yaw offset. The specific function was $f(\alpha_{pitch}, \alpha_{yaw})$ which performed best in the evaluation of Mayer et al. [28] to compensate the offset:

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

In the following, we will use the function in Equation (1) to investigate the following three questions: a) can we confirm the findings by Mayer et al. [27] that EFRC performs the best in VR; b) can we extend this for avatar independent models, in detail, we investigate if the correction using Equation (1)

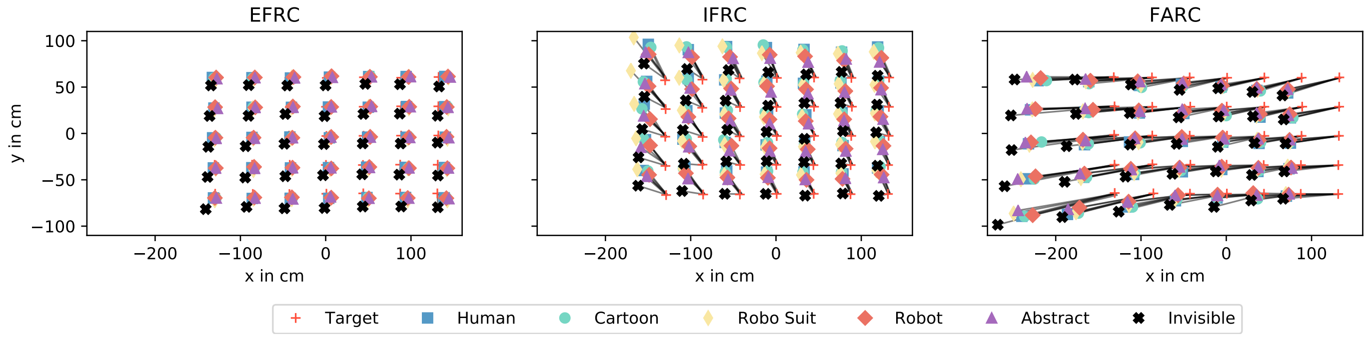


Figure 6. The displacement error for all ray casting methods (EFRC, IFRC, FARC). The colored points show the mean position of the ray intersecting with the virtual screen using each avatar.

is independent for AVATARS × RAY CAST; and c) does each avatar model need its own correction model or can we create a universal model?

To answer a) we built 18 models one for each AVATARS × RAY CAST and found that the average remaining error indeed is the lowest for EFRC, with 10.3*cm*, followed by IFRC (15.5*cm*) and FARC (32.3*cm*). Respectively this is an improvement of 3.88%, 57.38%, and 54.36%. The coefficients for the AVATAR dependent models are presented in Table 3.

In line with Mayer et al. [28], we validated these models using leave-one-participant-out cross-validation (LOPOCV) to ensure that we can reduce the offset using the function Equation (1) independent for each AVATARS, answering b). The results are shown in Table 2. Hence the function from Equation (1) can improve the pointing accuracy overall. Moreover, the results showed that the model could again improve EFRC on average by only 2.92%, while for IFRC we achieved a mean improvements of 43.93% and the best improvement was using the FARC with a offset reduction of 51.92%. However, the smallest remaining offset was achieved by using EFRC with an average offset of 10.59*cm* followed by IFRC with 20.71*cm* and FARC with 34.72*cm*.

Finally, to check if a universal model exits for all 6 avatars and thereby answering c), we built three models one for each RAY CAST technique. We used leave-one-out cross-validation (LOOCV) this time leaving *one avatar* out to access the quality of the models and if each AVATAR needs its own model. The smallest remaining offset was achieved using EFRC with an average offset of 10.85*cm* followed by IFRC with 15.84*cm* and FARC with 32.45*cm*. However, we found that this is only a offset reduction for IFRC and FARC (56.32% and 54.25% respectively) while the offset increased for EFRC by −1.0%. With respect to our 18 avatar dependent models, the three universal models performed worse for all RAY CAST techniques (EFRC: +5.24%, IFRC: +2.32%, and FARC: +.31%). Thus, a model for each AVATAR is beneficial to correct mid-air pointing errors.

DISCUSSION

We conducted a study in VR with 24 participants to investigate the effects of avatars on mid-air pointing accuracy. Our results show that the pointing accuracy of EFRC and IFRC was significantly affected by different avatars. Particularly, the *robot* and *abstract* avatars showed an increased accuracy of the IFRC. Inaccuracy does not correlate with the degree of perceived human-likeness. The increased accuracy of IFRC could be explained by the fact that the limb shapes of the *robot* and *abstract* avatar hands provide more visual information about the correct pointing direction of the index finger than a human, robot, or invisible hand. The error of the IFRC increases with the *cartoon* avatar, which can be explained by a lack of depth cues using the cel shader. We conclude that the error does not correlate with the degree of human-likeness. However, the rendering of virtual avatars decreases the pointing error in VR up to 45.3% (*human* vs. *invisible* avatar conditions) for EFRC as the most accurate ray casting method.

We can confirm the overall trend (cf. Foley et al. [13] in the real environment) that limbs tend to overreach targets to the opposite side of their favored arm (all participants were right-handed). We showed that the pointing error is systematically affected in horizontal as well as in vertical-direction, which means that both directions should be considered in determining or compensating the pointing error in VR. A polynomial correction model further increases accuracy by up to 57%. LOOCV ensured the validity of the presented models. We achieved the smallest remaining offset when using EFRC and a human-like appearance.

Interestingly, the FARC and IFRC methods in the *invisible* avatar condition did not show the expected increase of the mean error as measured in the EFRC method. It is clear that properly tracked and rendered avatar hands visually help users bring the position of the index finger into one line with the target and their eye. However, we assume that the forearm

	Remaining Offset (*cm*)			**Improvements (%)**		
AVATAR	EFRC	IFRC	FARC	EFRC	IFRC	FARC
Human	8.4	19.94	34.03	2.32	47.14	51.52
Cartoon	8.49	18.4	33.11	2.47	55.14	51.89
Robo Suit	9.62	20.7	33.94	.93	51.01	52.22
Robot	9.13	16.98	32.87	−.17	49.37	52.4
Abstract	9.78	15.95	34.74	.4	49.52	52.5
Invisible	18.11	32.3	39.64	11.57	11.41	50.98
Average	10.59	20.71	34.72	2.92	43.93	51.92

Table 2. Final LOOCV results using the correction Equation (1) proposed by Mayer et al. [28]. These results are archived using independent models for AVATAR × RAY CAST.

	Human		*Cartoon*		*Robo Suit*		*Robot*		*Abstract*		*Invisible*	
coeff	pitch	yaw	pitch	yaw	pitch	yaw	pitch	yaw	pitch	yaw	pitch	yaw
x_0	.003	.059	−.02	.07	−.018	.05	.002	.028	−.032	.041	−.016	.089
x_1	.098	.001	−.104	−.145	.789	−.622	.387	−.418	.345	−.389	−.098	−.024
x_2	.003	−.013	.006	−.03	.003	−.025	.002	−.021	.008	−.025	.048	−.048
x_3	.062	−.028	.011	−.06	−.01	.117	−.051	−.011	.051	−.029	.047	−.166
x_4	.216	−1.506	.384	−1.696	.995	−2.64	.335	−2.077	1.01	−2.455	2.187	−6.799
x_5	−4.082	2.991	−2.913	−.602	−8.341	−5.174	−5.452	−1.986	−8.522	−2.418	−38.352	5.016
x_6	−.025	.009	.004	.001	−.056	.042	−.029	.033	−.084	.075	−.031	.12
x_7	−.105	.608	−.388	.24	−.194	.897	−.53	.711	−.406	1.31	−1.088	1.165
x_8	.458	−1.146	−.694	−1.296	.244	−.554	.326	−1.331	.816	−2.112	2.763	−5.817
x_9	4.495	−45.881	37.221	−60.417	50.873	−46.149	14.571	−6.2	65.456	−49.404	24.503	−129.608
x_{10}	−39.805	32.051	35.683	94.242	−349.019	200.829	−162.74	143.21	−125.987	67.883	−314.623	−61.761
x_{11}	−22.595	22.705	−19.123	38.279	−11.299	−18.594	3.602	35.656	−32.53	34.765	15.507	117.665
x_{12}	−381.499	−1571.322	−483.672	−1678.815	−1266.141	−924.374	−425.234	−1369.7	−1348.018	−656.631	−2851.758	4304.068
x_{13}	−2622.557	−1483.611	−2934.119	−307.917	−1573.722	409.755	−2002.314	−590.455	−1034.562	−1323.71	4030.81	−3718.72
x_{14}	45160.606	−86356.916	35311.087	−91272.05	64306.238	−103109.811	37735.506	−113988.665	36316.381	−94136.501	290730.52	80019.157

Table 3. Model coefficients (in 10^{-5}) of the correction function using the IFRC method. The function is given by Equation (1). the six pointing devices (Human, Cartoon, Robo Suit, Robot, Abstract, Invisible Avatar). The coefficients are rounded with in the 95% confidence bounds.

and index finger are probably not affected to the same extent as the EFRC method, because of their overall unreliability while pointing. This implies that the orientation of finger and forearm does not have the influence while pointing we have suspected so far. This would also mean that users do not consciously pay attention to how forearm and fingers are oriented even when they see realistic bodies. Thus, our results do not only confirm previous work showing that the EFRC method is the most accurate method for determining the target but also that people primarily rely on the pointing direction between eye and finger tip and not on the directions of their forearm or index finger.

As we showed that the appearance of the VR avatar has a significant effect on the accuracy of the pointing gesture, the representation of the own body must be considered while applying compensating models in VR. The *human* avatar showed high accuracy in pointing and was perceived as the most human-like. None of the ray casting methods were significantly affected when only either texture or shading was changed (cf. *human*, *cartoon*, and *robo suit*). Potential effects or distraction of the human avatars regarding the uncanny valley or other phenomena could not be found in our study.

CONTRIBUTION

We contribute with our findings that avatars can affect the accuracy in pointing. We provide six mid-air compensation models for correcting the pointing offset using the most accurate ray casting method (EFRC) for six different avatar styles in virtual reality based on motion capturing data of 24 participants. These models enable finer selection tasks and precise interactions without using additional hand controllers or input devices as referential devices. This allows precise interaction with the VR system without additional controllers. Hand-free interaction increases the users' immersion and is an important step forward in further improvement of virtual applications and games. We further contribute with fully available source code and assets used in this experiment[3].

Design Implications

Assuming that people try to hold their finger, hand, and arm in an optimal manner by proprioceptive and visual information, we either recommend the use of realistic human avatars or human-like avatars as referential pointing device in VR or to use very abstract avatars where the body structure, geometry, and shading clearly indicate the pointing direction. For example, 2D shading such as used for cartoon rendered avatars, is not recommended when depth cues should provide reliable information about the correct pointing direction. For highest precision, we recommend to use EFRC. Considering the IFRC method, we recommend using less human-like virtual characters providing visual cues about the actual pointing direction of the index finger. To determine the correct direction of a users' pointing gesture, we recommend to use EFRC and the compensating model for the desired avatar as it significantly reduces the error in very immersive environments without additional visual feedback.

Limitations

More research is needed to understand the effects of VR avatars on referential gestures such as pointing. Even perfectly tracked limbs cannot avoid distractions and uncomfortable feelings caused by the virtual body representation. Though we increased the accuracy using polynomial models, we expect florian alt that the accuracy could be further improved, which is relevant for VR games and immersive applications. Furthermore, our participants have pointed at targets in front of them. We assume that there might be a difference when the target direction is different from the alignment of the body. We assume that including motion trajectories, additional body poses, and gaze directions in compensating further decrease the pointing error.

In the herein presented study, we used whole body avatars with a human-like body structure. Structural changes (*e.g.*, missing or additional limbs as investigated by Schwind et al. [39]) of avatars that can affect the pointing accuracy were not investigated. It is conceivable that there are more styles and morphologies of virtual avatars that potentially increase the accuracy in free-hand pointing, and it is important to understand to which extent an immersive VR experience still occurs using

---

[3] `https://github.com/interactionlab/pointing-in-vr-hands`

such virtual appearances of one's own body. Thus, a potential promising continuation of our work would be an examination of how the avatar models must be designed to reduce the error of the pointing gesture while providing high levels of presence in VR.

Future Work

To design new interaction techniques for VR and games using pointing gestures, it is important to understand which cues of the own avatar influence the accuracy in mid-air pointing. Future work should derive models explaining how informative cues in the virtual appearance can be used to further improve the interpretation of the correct pointing direction. Further research can use our models, for example, to improve locomotion through virtual worlds, for simple selecting tasks, or to improve the performance while interacting with virtual user interfaces. Other gestures beyond pointing can be considered and compensated when systematic errors are occurring. Previous work shows that eye gaze as ray casting technique can further increase the accuracy in pointing [20], however, requires additional equipment and calibration. Future work can explore a combination of our models and such supporting devices. Additional referential devices or indicators of the pointing direction such as cursors or rays potentially help to further increase the pointing accuracy in VR. Further research must explore how such devices can be designed that they do not negatively affect the VR experience. Finally, we also suggest to investigate if modeling and compensating the error in VR pointing could be further improved by rotating joints or body poses of the avatar.

ACKNOWLEDGEMENTS

This work was supported by the German Research Foundation (DFG) within project C04 of SFB/Transregio 161, the DFG Cluster of Excellence in Simulation Technology (EXC 310/2), and by the German Academic Exchange Service (DAAD) RISE Research Internship Program 2017.

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, 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`
3. Ferran Argelaguet, Ludovic Hoyet, Michaël Trico, and Anatole Lécuyer. 2016. The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *IEEE Virtual Reality (VR '16)*. Greenville, SC, USA, 3–10. DOI:`http://dx.doi.org/10.1109/VR.2016.7504682`
4. 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`
5. Bethesda Game Studios. 2017. *Fallout 4 VR*. Game [PC]. (December 2017). Bethesda Softworks, Rockville, ML, USA. Last played July 2018.
6. Bethesda Game Studios. 2018. *Skyrim VR*. Game [PC]. (3 April 2018). Bethesda Softworks, Rockville, ML, USA. Last played July 2018.
7. 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`
8. Sebastian Boring, Marko Jurmu, and Andreas Butz. 2009. Scroll, Tilt or Move It: Using Mobile Phones to Continuously Control Pointers on Large Public Displays. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7 (OZCHI '09)*. ACM, New York, NY, USA, 161–168. DOI: `http://dx.doi.org/10.1145/1738826.1738853`
9. 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`
10. 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`
11. Giambattista Della Porta. 1593. *De refractione Optices Parte: Libri Novem...* Ex officina Horatii Salviani, apud Jo. Jacobum Carlinum, & Antonium Pacem.
12. 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`
13. 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`
14. 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`

15. 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`

16. 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`

17. John B. Haviland. 2003. *How to point in Zinacantán*. Psychology Press, Mahwah, New Jersey, USA, 139–169.

18. Chin Chang Ho and Karl F. MacDorman. 2017. Measuring the Uncanny Valley Effect: Refinements to Indices for Perceived Humanness, Attractiveness, and Eeriness. *International Journal of Social Robotics* 9, 1 (jan 2017), 129–139. DOI: `http://dx.doi.org/10.1007/s12369-016-0380-9`

19. Hartwig Holzapfel, Kai Nickel, and Rainer Stiefelhagen. 2004. Implementation and Evaluation of a Constraint-based Multimodal Fusion System for Speech and 3D Pointing Gestures. In *Proceedings of the 6th International Conference on Multimodal Interfaces (ICMI '04)*. ACM, New York, NY, USA, 175–182. DOI: `http://dx.doi.org/10.1145/1027933.1027964`

20. 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`

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

22. Konstantina Kilteni, Ilias Bergstrom, and Mel Slater. 2013. Drumming in Immersive Virtual Reality : The Body Shapes the Way We Play. 19, 4 (2013), 597–605.

23. Pascal Knierim, Valentin Schwind, Anna Maria 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 (2018-04-21). *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (2018). DOI: `http://dx.doi.org/10.1145/3173574.3173919`

24. 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`

25. 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`

26. Jean-Luc Lugrin, Johanna Latt, and Marc Erich Latoschik. 2015. Avatar anthropomorphism and illusion of body ownership in VR. In *2015 IEEE Virtual Reality (VR)*, Masataka Imura, Pablo Figueroa, and Betty Mohler (Eds.). IEEE, 229–230. DOI: `http://dx.doi.org/10.1109/VR.2015.7223379`

27. Sven Mayer, Valentin Schwind, Robin Schweigert, and Niels Henze. 2018. The Effect of Offset Correction and Cursor on Mid-Air Pointing in Real and Virtual Environments, In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (2018-04-21). *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (2018). DOI: `http://dx.doi.org/10.1145/3173574.3174227`

28. 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`

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

30. 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`

31. 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`

32. Masahiro Mori, Karl F. MacDorman, and Norri Kageki. 1970/2012. The Uncanny Valley [From the Field]. *Robotics & Automation Magazine, IEEE* 19, 2 (1970/2012), 98–100. DOI: `http://dx.doi.org/10.1109/MRA.2012.2192811`

33. 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`

34. Brad A. Myers, Rishi Bhatnagar, Jeffrey Nichols, Choon Hong Peck, Dave Kong, Robert Miller, and A. Chris Long. 2002. Interacting at a Distance: Measuring the Performance of Laser Pointers and Other

Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '02)*. ACM, New York, NY, USA, 33–40. DOI: http://dx.doi.org/10.1145/503376.503383

35. 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

36. 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

37. T. Pfeiffer, M. E. Latoschik, and I. Wachsmuth. 2008. Conversational Pointing Gestures for Virtual Reality Interaction: Implications from an Empirical Study. In *Proceedings of the 2008 IEEE Virtual Reality Conference (IEEE/VR'08)*. 281–282. DOI: http://dx.doi.org/10.1109/VR.2008.4480801

38. 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

39. Valentin Schwind, Pascal Knierim, Lews Chuang, and Niels Henze. 2017a. "Where's Pinky?": The Effects of a Reduced Number of Fingers in Virtual Reality. In *Proceedings of the 2017 CHI Conference on Computer-Human Interaction in Play (CHI PLAY'17)*. ACM, New York, NY, USA, 6. DOI: http://dx.doi.org/10.1145/3116595.3116596

40. Valentin Schwind, Pascal Knierim, Cagri Tasci, Patrick Franczak, Nico Haas, and Niels Henze. 2017b. "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

41. Valentin Schwind, Katharina Leicht, Solveigh Jäger, Katrin Wolf, and Niels Henze. 2017c. Is there an Uncanny Valley of Virtual Animals? A Quantitative and Qualitative Investigation. *International Journal of Human-Computer Studies* 111C (2017), 49–61. DOI: http://dx.doi.org/10.1016/j.ijhcs.2017.11.003

42. Valentin Schwind, Lorraine Lin, Massimiliano Di Luca, Sophie Jörg, and James Hillis. 2018. Touch with Foreign Hands: The Effect of Virtual Hand Appearance on Visual-Haptic Integration. In *Proceedings of the ACM Symposium on Applied Perception (SAP '18)* (2018-08-03). ACM, New York, NY. DOI: http://dx.doi.org/10.1145/3225153.3225158

43. 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

44. Mel Slater, Martin Usoh, and Anthony Steed. 1995. Taking Steps: The Influence of a Walking Technique on Presence in Virtual Reality. *ACM Transactions on Computer-Human Interaction* 2, 3 (Sept. 1995), 201–219. DOI:http://dx.doi.org/10.1145/210079.210084

45. Vertigo Games and Jaywalkers Interactive. 2016. *Arizona Sunshine*. Game [PC,PS4]. (6 Decembter 2016). Vertigo Games, Odessa, TX, USA. Last played July 2018.

46. Vinoba Vinayagamoorthy, Andrea Brogni, Marco Gillies, Mel Slater, and Anthony Steed. 2004. An investigation of presence response across variations in visual realism. In *The 7th Annual International Presence Workshop*. 148–155.

47. 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

48. Juan Pablo Wachs, Mathias Kölsch, Helman Stern, and Yael Edan. 2011. Vision-based Hand-gesture Applications. *Commun. ACM* 54, 2 (Feb. 2011), 60–71. DOI:http://dx.doi.org/10.1145/1897816.1897838