The Effects of Gender and the Presence of Third-Party Humans on Telepresence Camera Height Preferences

Kevin Pfeil
University of Central Florida
kevin.pfeil@knights.ucf.edu

Pamela J. Wisniewski
University of Central Florida
pamwis@ucf.edu

Joseph J. LaViola Jr.
University of Central Florida
jjl@cs.ucf.edu

ABSTRACT

With the rise of immersive virtual reality and telepresence, it is important to understand the factors that contribute to creating an optimal user experience. In particular, there are divergent recommendations for setting camera heights in virtual contexts that facilitate telepresence. Therefore, we conducted a 2x2 mixed-design experiment with 93 college students asking them to select their preferred camera height when varying camera placement (overhead, chest) and the presence of human avatars (present, not present). We found that while camera placement did not have a significant effect of preferred camera height, the presence of avatars (increased height preference) and gender (women preferred lower heights) were significant. Our results provide evidence that factors within a virtual environment and individual differences influence users' preferences of camera height. Thus, systems designed for immersive virtual reality and telepresence should customize camera height based on these factors.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; **User studies**.

KEYWORDS

telepresence, virtual reality, camera height, user study

ACM Reference Format:
Kevin Pfeil, Pamela J. Wisniewski, and Joseph J. LaViola Jr.. 2020. The Effects of Gender and the Presence of Third-Party Humans on Telepresence Camera Height Preferences. In *ACM Symposium on Applied Perception 2020 (SAP '20), September 12–13, 2020, Virtual Event, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3385955.3407924

1 INTRODUCTION

Immersive Virtual Reality (VR) is making a resurgence with the release of more affordable, consumer-grade hardware such as the HTC Vive head-mounted display (HMD); the advent of 360° / panoramic cameras such as the Ricoh Theta V; and the ease of access to content hosting platforms such as YouTube. With these tools, even novice users are able to create immersive content and share these experiences with others. For instance, there is a recent trend in which people (or "streamers") wear cameras on their body, and live-stream their environment while walking around their physical surroundings. This setup allows a viewer to receive an immersive tour of any given remote location that has a streamer within it. As such, this interaction is a form of telepresence, which is the ability to perceive and/or interact with a remote environment, as if being there [26]. This kind of telepresence has been applied for activities such as shopping [5], running [1], and sightseeing [6], among others.

A primary objective of immersive VR and telepresence is to make the experience as natural as possible, as if the viewer was actually present in the remote or virtual environment. One factor that can greatly influence the users' experience is the point-of-view in which the viewer interacts; depending on how tall the camera is, the view may or may not be optimal for the audience. It is thus increasingly important to understand what factors contribute to an enjoyable experience for all. As such, previous efforts have studied optimal camera placement on the body (e.g., overhead, shoulder, or chest) [33], and have found that those which have the most unoccluded view (e.g. overhead rather than on the shoulder) accommodate the most opportunity for exploration, and are thus preferred most.

Meanwhile, others have studied the effect of third-party avatars on distance perception within a given virtual environment (VE). Some have found that these avatars do not have an effect [2, 32], but Langbehn et al. revealed that they help to provide a sense of scale within a VE [21]. Given the importance of camera height, camera placement, and the presence of human avatars within a VE, we aim to understand how these factors influence one another. Additionally, we are interested in how individual differences, such as a users' height and gender, play a role in their optimal camera height preference. Previous efforts suggest that camera height is relatively arbitrary, and that the optimal height for all users is approximately 150cm [16, 41]. Building upon these previous works, in this paper we allow users to manually set the camera height in various conditions. We ask the following research questions:

- **RQ1**: *What factors (camera placement and avatar presence) are important for determining optimal camera height based on user preference?*
- **RQ2**: *What individual differences (gender, user height) play a role in determining user preference of camera height?*
- **RQ3**: *What is the acceptable range (shortest, preferred, tallest) of user-selected camera heights based on the factors and individual differences above?*

To answer these research questions, we conducted a 2x2 mixed-design experiment with 93 participants, with a within-subjects factor of Camera Placement, and a between-subjects factor of Avatar Presence. While viewing an immersive VE that simulated a telepresence experience with a body-worn 360° camera, participants were asked to manually set the height of the camera according to what they felt was most natural to them. Additionally, we asked them to identify

SAP '20, September 12–13, 2020, Virtual Event, USA

ACM ISBN 978-1-4503-7618-1/20/07…$15.00
https://doi.org/10.1145/3385955.3407924

their shortest and tallest preferred camera heights. Results show that Camera Placement was not a significant factor that influenced height preference, but human avatar presence was significant; when avatars were present, participants preferred taller camera heights. The average camera height preference for men when avatars were present was around 160cm; but, women preferred a shorter height - approximately 140cm. When avatars were not present, average camera height preferences were approximately 145cm and 120cm, respectively. Actual user height was not found to play a role in height preferences. Lastly, the average thresholds for acceptable height were approximately +/- 40cm from a user's natural height selections.

Our work contributes a user-centric approach to understanding camera height preferences within an immersive VE, with a call to consider individual differences instead of choosing a "one-size-fits-all" solution. Though our research agenda targets optimal camera height parameters for telepresence viewers, and though there is strong conceptual overlap between telepresence and virtual reality, we acknowledge that we cannot directly generalize our results to the real world. Since the use of a VE allows us to keep external variables constant, we elect to conduct our research in this simulation, while being cautious to make direct claims for the real world. Our main contributions include the following:

- Our work is the first to examine user preference of camera height relative to camera placement and the presence of avatars.
- We found that the virtual context (presence of human avatars), rather than the physical placement of the camera on the host, had a significant influence on users' preferred camera height.
- We also highlight the importance of taking into account individual differences, specifically gender, when customizing virtual telepresence user experiences to meet users' needs.

2 RELATED WORK

In this section, we first synthesize the literature that demonstrates why camera height is an important factor to consider when studying immersive virtual telepresence and its relation to camera placement. Next, we highlight how judging distances within VR may affect preferred camera height and how the presence of human avatars may help users more accurately judge height.

2.1 Immersive Telepresence and Virtual Reality

Considerable work has been done at the intersection of VR and telepresence, and the two are strongly linked. Regarding traditional telepresence, Minsky hypothesized that telecommunications could allow a worker to perform duties over long distances through use of a robotic platform linked to telecommunications [26]. Over time, this definition has broadened to include any interaction that facilitates the perception and/or manipulation of a remote environment, as if actually there. For instance, telepresence includes popular video-chat applications such as Zoom and Skype, which are typically used on flat screens such as laptops and smartphones. However, the user can experience a more *immersive* interaction with the use of a VR HMD. By wearing an HMD, a user blocks out the "real" world, and is enveloped in the one they are viewing [46, 52].

VR HMDs are typically used to provide users with a way to explore a virtual world from a first-person point of view; as they move their head, the viewpoint changes intuitively [23]. VEs are often created with computer graphics, but they are not always fictional. With the release of 360° imaging hardware such as the Ricoh Theta V, any real environment can be captured and replicated as a virtual world. With proper equipment, a remote environment can be captured, digitized, transmitted, and reconstructed in VR, in real-time - and in this way, an immersive telepresence experience can be shared with someone across the globe.

As such, VR and telepresence heavily overlap. VR has been used to simulate real-life experiences for a variety of use cases, including surgery [44], flight [53], warfare [4], robotics [7], and even Augmented Reality [36]. In this paper, we use VR to simulate a simple immersive telepresence experience. As both use egocentric points of view, and as VR allows us to control external variables, we find VR to be natural testbed for our research goals.

2.2 Immersive Telepresence Camera Placement

A common feature of virtual reality films generated with 360° cameras is the inclusion of a static, unmoving tripod to capture surroundings. This allows a user to explore a particular view without any occlusion, and prevents simulator sickness caused by a disagreement between the visual and vestibular systems [15]. Though effective, the video is typically constrained to previously-recorded footage. For a more personal experience, people can live-stream a video feed to another while moving about an environment. To free their hands during this interaction, streamers tend to wear a camera somewhere on their body.

A variety of camera placements have emerged in the body-worn telepresence community, including on or above the head [10–12, 27, 28, 48], on the shoulder [13, 18–20, 43, 51], and on the chest [3, 9, 50], among others. For a given streamer, though, placing the camera in these locations affects the view height, and it also changes the quality of view. From a recent study, it is recommended to place the camera on a bodily position that will offer the viewer maximal exploration capabilities, e.g. over the streamer's head [33]. However, it is not yet clear how the height of the streamer (and thus the resulting height of the camera) affects user experience.

Other researchers have studied the importance of camera height directly, and a noticeable trend is the recommendation of a "one-size-fits-all" value, not accounting for individual differences of users. Many divergent recommendations have been issued for static shots or pictures, such as keeping the camera "eye-level with your subject" [47], "at chest-level" [22], or at "person-height" [35]. Keskinen et al. found that optimal camera height is around 150cm [16], and Rothe et al. suggested a height of 156cm if the user is standing [41], but these were for static, unmoving cameras.

Building upon this prior work, we aim to understand optimal camera height from a user-centric point of view. Since telepresence with a human streamer affords a moving viewpoint, we conducted a study in which the camera is constantly moving through a visually rich environment. As we are interested in user preference, we had our participants freely manipulate and select heights in real time. Additionally, we measured their Tallest and Shortest height preferences in order to understand if there is an acceptable range of heights that would suffice for an immersive telepresence experience. Since it is very difficult for a human streamer to wear a camera *and* have it be at a specific height, it is important to know how much deviation from a "natural" height is acceptable.

2.3 Distance Estimation in Virtual Reality

In order to accurately judge camera height in a virtual environment, one must be able to estimate distance. The virtual reality community has perennially studied human ability to perceive and estimate distances in VEs, and a common theme from the emerging literature is that humans tend to underestimate distances while wearing an HMD [29, 39]. Since immersive telepresence viewers may use VR equipment, we suspect that they will be subject to this problem, and it will manifest itself in the form of height underestimation.

There are a variety of factors which affect distance perception; for instance, there are some which are specific to the VR hardware. It has been shown that higher graphical fidelity increases accuracy [34, 49], and manipulating interpupillary distance (IPD) of an HMD affects judgments [17]. A recent work suggests that with technological advances in HMDs (e.g. better screen resolution), the perceptual gap will shrink [14]. Thus, over time, better HMDs will help to reduce this difference. However, hardware characteristics only explain a piece of the problem. It has been shown that environmental contexts can also affect distance perception. Some work has found that, in VR, human self-avatars (first-person characters that represent a user) give a frame of reference and thus increase accuracy [24, 30]. Third party human avatars may also affect distance perception; some research suggest that they do not help alleviate this issue [2, 32], but others have shown that scale is interpreted differently depending on their presence [21, 37]. Though scale and distance are not quite the same, they are related; our work in part helps to determine the effect of human avatars on height judgment calls.

We acknowledge that users are generally less adept at estimating distances while wearing a VR HMD, and we suspect that this phenomenon will translate to immersive telepresence experiences. We thus recall these previous findings and, as appropriate, eliminate them as variables in our study. We use a high-resolution HMD (HTC Vive) and change IPD as needed per participant; and, as a self-avatar is similar to a telepresence streamer, we keep one visible at all times. We also manipulate third-party avatar presence as an independent variable to help understand how camera height preferences shift with varying social environments. Since various contexts can have an effect on optimal camera heights for telepresence, it is pertinent to understand the influence third party members have on an experience.

3 METHODS

Previous research efforts have shown how various factors impact user experience, including camera placement [33], and the presence of human avatars [2, 21, 24]. We based our study on these factors and applied a user-centric approach by letting participants identify their own preferred camera height. On demand, our participants could increase or decrease the height of the camera in a simulated telepresence experience.

3.1 Study Design

We conducted a 2x2 mixed-design experiment with independent variables (IVs) of *Camera Placement* and *Avatar Presence*. The *Camera Placement* variable was a within-subjects factor and had 2 levels based on previous work - "Overhead," where the camera is mounted above the avatar's head, and "Chest," where the camera is mounted on the avatar's chest [33]. Both were first-person views, i.e. the user could rotate the view by simply moving their head. The *Avatar Presence* variable had 2 levels - Present and Absent - and was between-subjects to prevent a learning effect. In the Avatars Present condition, all third-party human avatars were visible. In the Absent condition, they were invisible, with the exception of the streamer. We used this variable because we anticipated participants making height judgment calls using human avatars as a frame of reference, as per Langbehn et al. [21]. We randomly assigned participants to their corresponding group and condition order, in a counter-balanced design. Our study received IRB approval.

3.2 Using Virtual Reality to Simulate Immersive Telepresence

Our study is comparable to those of Keskinen et al. [16] and Rothe et al. [41], who used real 360° footage from different heights (approx. 10cm steps) from a tripod-mounted camera. Their research was related to static 360° video, and thus they used real, prerecorded footage. Our research interests focus on mobile telepresence, and here we simulate a live-streamed 360° camera using a VE, due to the many external and possibly confounding factors which can hinder our measurements.

First, as we are interested in telepresence, we must consider how the camera is worn by a moving person instead of static on a tripod. With a single actor, it is nigh impossible to maintain a constant walking gait for a number of video clips (i.e., people have different walking motions [42], which could confound a study). Second, we must also consider the height of the streamer. In order to preserve camera placement while manipulating height, we would need to enlist multiple actors of incremental height, which confounds visual identity of the streamer in addition to exacerbating the walking gait problem. We considered alternatives to help alleviate this issue - one idea was the development of an actuated, mobile platform on which the same person could stand (thus changing their height while preserving visual identity); another was shrouding actors with a sheet, to hide their faces and bodies. Both of these do not completely solve the above issues, and perhaps worse still, depart from any semblance of telepresence. We thus decided to simulate a 360° camera in a VR setting, allowing us to reasonably overcome these aforementioned problems.

3.3 Apparatus

We used a "Showroom Environment" found on the Unity Asset Store [45], as a template to build our VE. The showroom comes with hangers for paintings and low pedestals for sculptures. Sight-seeing tours (real or virtual) are common VR experiences [25, 48], so we used the environment to build a fictitious one. See Figure 1 for an illustration of the environment.

On the hangers, we placed images of famous paintings, such as Mona Lisa, The Starry Night, and The Kiss, finding a mixture of small, medium, and large paintings. We found the exact dimensions for each painting on the internet, and used these values to scale the paintings so that they were life-sized. The center of each painting was at a height of 177cm measured from the floor. In total we used 15 famous paintings. On the pedestals, we placed a variety of everyday objects which we believed would provide familiarity and thus a frame of reference for size. For instance, we added a can of Pepsi, a soccer ball, and a Mac computer. The pedestals were each approximately 69cm tall, measured from the floor, and the sculptures were of varying size. In total, we used 12 sculptures on the pedestals. On the floor, we laid out ornate rugs to provide more interesting stimuli below eye level, using various textures found online. In total, there were 6 ornate rugs.

In addition to these stimuli, our environment had two male and two female human avatars. Both male avatars were approximately 175cm tall, and both females were approximately 163cm tall, following average height in the United States [8]. Third-party avatar height was kept constant for all participants in our study. These avatars did not move except for a simple "breathing" animation. All stimuli were mapped symmetrically in the environment. In our environment, a virtual human (the Ethan model from Unity3D tutorials) provided a social telepresence experience to our users. This simulated streamer walked the scene at a constant speed, in a box-shaped path around the room. In this way, we were able to keep navigation constant in our study. If the user looked around, they would be able to see the streamer's face and body.

There are numerous 360° cameras available for professional and consumer purchase, and we created a 6-camera rig in Unity3D. This, combined with the 3D effect afforded by VR, is not unlike the Vuze+ stereo 3D camera[1]. Further, we applied the following Unity3D post-processing techniques (to be consistent with the study by Pfeil et al. [33]), resulting in the common visual artifacts seen when live streaming multi-lens camera feeds:

- **Antialiasing** (Fast Approximate Anti-Aliasing)
- **Ambient Occlusion** (Intensity = 1; Radius = 0.3; Sample Count = Medium; Downsampling Enabled)
- **Motion Blur** (Shutter Angle = 270; Sample Count = 10; Frame Blending = 0)
- **Grain** (Intensity = 0.5; Luminance Contribution = 0.8; Size = 0.7; Colored Enabled)

We used the 2017.3.0f3 version of Unity3D to display our VE. It was displayed using a VR-ready laptop with Windows 10, Intel core i7-7700HQ at 2.8GHz, with 12GB of RAM, with an Nvidia GeForce GTX 1060. The HTC Vive was used for our study, and the users were given an HTC Vive controller to provide inputs as appropriate.

3.4 Procedure

Recruited participants were seated inside of the HTC Vive play area. We collected demographics and explained the purpose of the study. We noted the height of the participant and subtracted the difference between their eyes and the top of their head (approx. 15cm) to record their Eye Height. Following, we explained that their objective was to manipulate the height until it felt the most natural to them, and to identify how tall and how short the camera could go, before it negatively affected their experience.

[1] https://vuze.camera/camera/vuze-plus-camera

Figure 1: Sample snapshot of the virtual art museum. Stimuli were placed symmetrically throughout the room, and the users' simulated telepresence streamer followed a box-shaped path between these stimuli.

We randomized and logged the initial camera height; the HTC Vive play area was minimized such that participants' head motions (aside from rotation) would not cause a change in camera placement. Thus, user height did not cause a change in display. Using the random initial height, we correspondingly scaled the virtual streamer, such that the camera remained in the same place on its body during the study. For purposes of our study, we assumed camera stabilization; i.e., the streamer's walking gait had no effect on the camera, except for when the avatar turned a corner. Inside the VE, the streamer walked around the museum one full rotation before we asked any questions, to let users familiarize themselves with the environment. After the streamer completed the round, we gave the participant an HTC Vive controller. By clicking on the top or bottom of the thumb pad, they could increase or decrease the camera height by 2.54cm (1 inch), respectively. The virtual streamer scaled with the camera, so to be visually consistent. When the user increased the camera height, the streamer became taller. When the user decreased the height, the streamer became shorter. The streamer continued walking in the museum at a constant speed, regardless of scale.

For the participants in the Avatars Absent conditions, we removed the male and female avatars completely from their view, except for the simulated streamer. For the participants in the Avatars Present conditions, these were visible to our participants. The height for each of these third-party avatars remained constant across all runs. We did not inform the participants how tall these avatars were, instead drawing out their responses through their own perception of the environment.

We then asked the user to adjust the camera until it arrives at the most natural height for them; we logged this value. Next, we asked them to slowly increase the camera height until the view started to become too tall; we logged this value. We then manually reset the camera height to the value corresponding to the "most natural" height, and then asked the participant to slowly lower the view until it started to become too short; we logged this value. After collecting the data, we reset the environment and toggled the *Camera Placement* to the next condition, repeating these steps. The time to complete the

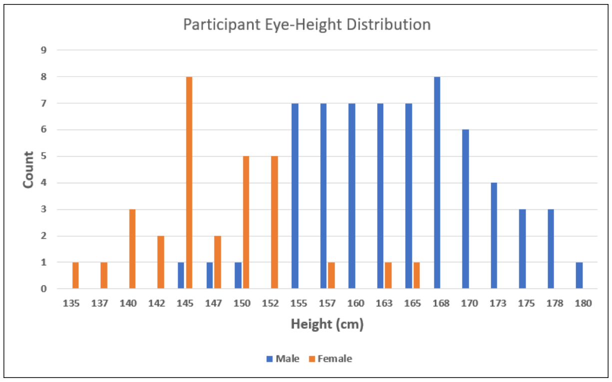

Figure 2: Histogram of participant eye-height, by Gender.

study was approximately 15 minutes, and participants were given 5 USD in cash.

3.5 Research Hypotheses

There are a number of factors which can play a role in determining a user's preferred camera height. Pfeil et al. found that different camera placements on a streamer's body affect user perception of height [33]; multiple researchers found that third-party avatars help to provide a sense of scale [21, 37] in a VE; and, expectedly, camera height preferences should scale with a user's actual height. As such, we hypothesize the following:

- **H1**: Participants will prefer significantly lower camera heights when the camera is placed overhead.
- **H2**: Participants will prefer significantly lower camera heights when avatars are not present.
- **H3**: A participants height will be significantly associated with their preferred camera height.
- **H4**: Females will prefer a significantly lower camera height than males.

3.6 Dependent Variables and Data Analysis Approach

As noted in our procedure, we collected data points from each user per condition. These data points were how tall the camera was in the environment, in centimeters. Per condition, we measured the user's height selections for *Natural*, *Tallest*, and *Shortest*. We used a Shapiro-Wilks test of normality to determine if our dependent variables were normally distributed, and they were (Natural: p = .099; Tallest: p = .186; Shortest: p = .139). To address our research hypotheses, we conducted an analysis of covariance (ANCOVA), with *Camera Placement* and *Avatar Presence* as fixed factors, and participant height and gender as covariates. Additionally, we use descriptive statistics in order to detail the acceptable range of camera heights.

3.7 Subjects

We recruited 97 college students from the University of Central Florida using an email blast, but we identified 4 who gave us outlier data points (values outside of three standard deviations); we removed them. Our resulting participant pool consisted of 63 males and 30 females. Their age range was from 18 to 44 (M = 22.2; SD = 4.17). We measured participant eye-height; the range was 135cm to 180cm (M = 159cm; SD = 10.7cm), following a normal distribution; see Figure 2 for an illustration of participant height. All participants had normal vision, or they wore corrective lenses during the study. On a 5pt scale (1 = Never, 5 = Very Frequent), the mean VR experience among our users was 2.1 (SD = 1.3).

Camera Placement	Avatar Presence	Gender	Preferred Camera Height
Chest	Absent	M	M = 142.7, SD = 30.23
		F	M = 118.2, SD = 32.26
	Present	M	M = 161.0, SD = 17.91
		F	M = 140.7, SD = 24.41
Overhead	Absent	M	M = 147.9, SD = 29.82
		F	M = 117.5, SD = 28.72
	Present	M	M = 166.1, SD = 20.49
		F	M = 147.0, SD = 20.02

Table 1: Descriptive Statistics of Preferred Camera Height (cm) by Camera Placement, Avatar Presence, and Gender

Camera Placement	Gender	Preferred Camera Height
Chest	M	M = 145.1, SD = 29.30
	F	M = 129.5, SD = 30.35
Overhead	M	M = 157.8, SD = 26.63
	F	M = 132.2, SD = 28.58

Avatar Presence	Gender	Preferred Camera Height
Absent	M	M = 145.3, SD = 29.88
	F	M = 117.9, SD = 30.02
Present	M	M = 163.6, SD = 19.27
	F	M = 143.8, SD = 22.17

Table 2: Descriptive Statistics of Preferred Camera Height (cm) by Camera Placement and Gender, and Avatar Presence and Gender

4 RESULTS

We first present the descriptive characteristics of our data, followed by hypothesis testing using a repeated-measures ANCOVA.

4.1 Descriptive Statistics

Tables 1 and 2 describe the mean heights and standard deviations for our dependent variables (DVs). On average, when human avatars were absent, participants set the height of the camera closer to the floor. Further, the participant responses for preferred camera height has far less variance when avatars were present in the scene; see Figure 3. Additionally and expectedly, men set the camera heights taller than women. Further, our participants mainly set the camera only slightly higher in the Overhead conditions.

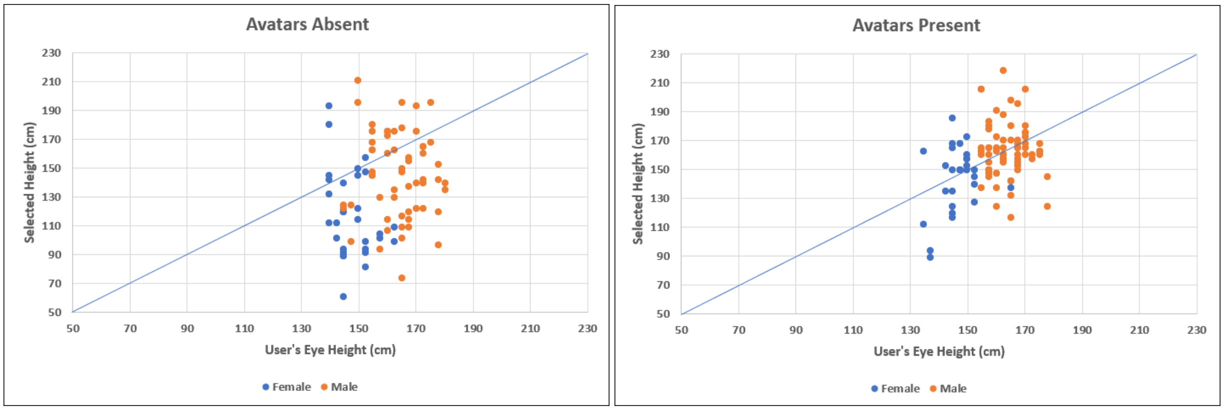

Figure 3: Scatterplots of user preferred camera height to actual eye height, by Gender. Left: Avatars Absent conditions. Right: Avatars Present Condition. Blue lines illustrates 1:1 ratio. In the presence of avatars, our participants clustered their responses closer towards their actual height than when avatars were absent.

Effect on Natural Height	ANCOVA Result
Main Effects	
CP	$F(1,89) = 0.541, p = .464, \eta_p^2 = .006$
AP	$F(1,89) = 18.12, p < .001, \eta_p^2 = .169$
UG	$F(1,89) = 12.48, p < .001, \eta_p^2 = .123$
UH	$F(1,89) = 0.331, p = .567, \eta_p^2 = .004$
Interaction Effects	
CP * AP	$F(1,89) = 0.269, p = .606, \eta_p^2 = .003$
CP * UH	$F(1,89) = 0.391, p = .533, \eta_p^2 = .004$
CP * UG	$F(1,89) = 0.664, p = .417, \eta_p^2 = .007$

Table 3: Results from a Repeated Measures ANCOVA regarding effects of Camera Placement (CP), Avatar Presence (AP), User Gender (UG), and User Height (UH) on Natural Height Preference

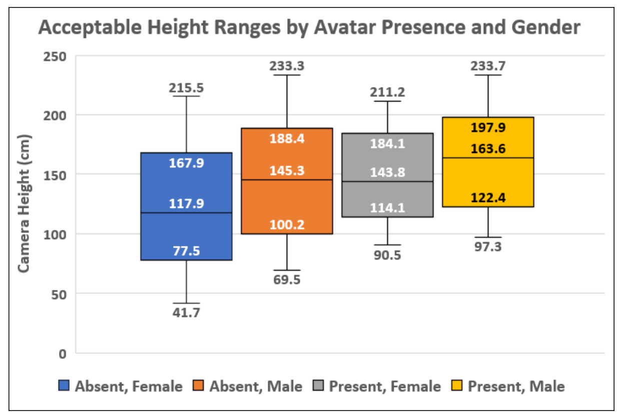

Figure 4: Average range of acceptable camera heights, by avatar presence and user gender. There is a very wide buffer surrounding the average camera height preferences which still affords an acceptable view. Black bars represent 1 σ.

4.2 Hypothesis Testing Results (RQ1 & RQ2)

We performed hypothesis testing using a repeated measures ANCOVA; Table 3 summarizes our statistical findings. *Camera Placement* was a within-subjects variable, and *Avatar Presence* was a between-subjects variable. Participants' gender and eye height were treated as covariates.

To test H1, which was that participants would compensate for seeing a streamer under them in the Overhead condition, we tested the main effect of *Camera Placement*. We did not find a significant main effect of *Camera Placement* ($F(1,89) = 0.541, p = .464, \eta_p^2 = .006$), which means that users' preference of camera height did not vary based on the camera being placed overhead or on the streamer's chest.

To test H2, which was that third-party avatars would provide a sense of scale and thus lead to users selecting a camera height closer to their own height, we tested the main effect of *Avatar Presence*. We found a significant main effect ($F(1,89) = 18.12, p < .001, \eta_p^2 = .169$), such that the absence of avatars lead to users setting the camera height lower to the ground. This also indicates that the presence of avatars helps users to select heights closer to their own.

To test H3, which was that participants would prefer a camera height that is closer to their own, we tested the main effect of User Height. We did not find a significant main effect ($F(1,89) = 0.331, p = .567, \eta_p^2 = .004$), which indicates that a user's height did not directly influence their camera height selections.

To test H4, which was that females would prefer a lower camera height than males, we tested the main effect of Gender. We found a significant main effect ($F(1,89) = 0.331, p = .567, \eta_p^2 = .004$), such that women did select lower camera heights than men.

We did not find any significant interactions among our variables.

4.3 Analysis of Height Preference Thresholds (RQ3)

To address our RQ3, we analyzed the responses given by our participants for their Tallest and Shortest heights. This assists us in determining the average range of acceptable camera heights; see Figure 4 for illustration. When avatars were absent, the average Tallest height thresholds were 167.9cm for females (SD = 47.6cm), and 188.4cm for males (SD = 44.9cm), and the average Shortest height thresholds were 77.5cm (SD = 35.8cm) and 100.2cm (SD = 30.7cm), respectively. This results in an average range of acceptable heights that is approximately 90cm wide for both men and women.

When avatars were present in the scene, the average Tallest height thresholds were 184.1cm for females (SD = 27.1cm), and 197.9cm for males (SD = 35.8cm), and the average Shortest height thresholds were 114.1cm (SD = 23.7cm) and 122.4cm (SD = 25.1cm), respectively. This results in an average range of acceptable heights that is approximately 72cm wide for both men and women.

Both of these results indicate that the acceptable range of camera heights is *very wide*. These ranges completely encompass average human height in the United States minus 3 standard deviations [8]; but, only in the Avatars Present conditions do we find the range to encompass average human height plus 2 standard deviations. This means that any given viewer should have an agreeable user experience, if the streamer wears the camera at their eye level or below.

5 DISCUSSION

Our results offer various insights into optimizing camera height for user experiences. In this section, we describe our recommendations for telepresence designers, which include considerations for individual differences and various social environments.

5.1 Contextual Cues within the Environment Matter More Than Camera Placement

Our results indicate that avatars in view of the camera provide a significant frame of reference, strongly affecting a user's preferred camera height selection (H2). It is important to consider the social landscape when conducting an immersive experience. If it is known to have people walking around (e.g. a popular art museum, concert, or sports venue), it becomes more appropriate to set the camera height closer to the viewer's natural height. In a socially empty scene (e.g. nature trail), camera height becomes more arbitrary. Additionally, our study uses a novel approach in the light of VR distance underestimation, and though our task was unique compared to previous efforts, our results corroborate with prior findings that avatars indeed provide a strong frame of reference [21, 37].

Per work by Aseeri et al., changes in the scale of third-party avatars do not seem to have an effect on distance estimation. They created a life-like 3D model of one of the researchers for their study and found that their measurements were unaffected when comparing the normal scale of the model to an increased or decreased scale (+/- 20%) [2]. An interesting contrast between our works is that their participants met the researcher in-person prior to the study (and thus had a true frame of reference for the 3D model), whereas in our study, our participants were not given any a priori knowledge of the avatars - they used their intuition to choose camera height in the presence of these avatars. Yet, by simply seeing these avatars in the scene, our participants were able to cluster their responses closer to their own height. The 20% buffer as described by Aseeri et al. is quite large; for our selected avatars which were scaled to average human height in the United States, a 20% difference would encompass more than 3 standard deviations from the mean [8]. This suggests that human avatar height is somewhat arbitrary when designing a VE, at least when considering the problem of distance underestimation. Assuming these findings translate to real-world telepresence cases, we would not expect any drop-off in user experience in socially-crowded scenes which are diverse in human height. The main consideration, then, is the presence vs. absence of third-party humans.

We also found that bodily camera placement on a streamer is not a significant factor that influences user height selections. This contradicts prior work in which an Overhead camera placement influenced participants to believe the camera was taller than it actually was [33]. Our intuition tells us that this finding stems from a difference of study procedure. In the previous study by Pfeil et al., participants were shown three unique camera heights (users' natural height and +/- 30cm) from the perspective of three bodily placements [33]; in our study, participants had the ability to change height at will, at 2.54cm increments. By allowing participants to explore heights in-depth, we expect that our study elicited more accurate personal preferences.

5.2 Gender Differences Matter

Although we anticipated our user-centered procedure to draw out user height as a significant factor (H3), our results corroborate with previous literature in that it is not [16]. We did find that user gender is a significant factor to consider when developing an immersive experience; this does not seem to be due to psychological differences, but rather due to natural physiological differences between the sexes which manifested in our participant sample. Expectedly, men chose heights approximately 20cm taller than women prefer (H4), and while this is a seemingly obvious finding (as height is correlated with gender), to our knowledge, no recommendations in prior literature for optimizing camera height have considered this gender difference. Our results show that, when possible, VR experiences could be tuned to the different sexes. In both academic literature and online articles pertaining to 360° film, it is currently suggested to set the camera at a static height [16, 22, 35, 41], but our findings indicate females would want a different view compared to males, to provide a more natural feeling. As such, we recommend immersive experience designers to consider this difference, to provide optimal user experiences for all, instead of a one-size-fits-all approach.

5.3 Identifying Appropriate Telepresence Streamers

Our study points to a large buffer of acceptable camera heights for telepresence. This means that for any given user viewing a camera feed in VR, there is a wide range of acceptable streamers who can provide a positive viewing experience; and, since camera placement does not seem to influence preferred view height, a given streamer can adjust a body-worn camera as needed, without worrying about negatively affecting the view quality. The main consideration, then, is the ability to freely explore the environment unoccluded. In general telepresence use, we are confident that average human height will not detract from the experience. However, we acknowledge that there are prominent use cases in which the streamers will push past these thresholds. For instance, the National Basketball Association (NBA) has been ramping up court-side live-streaming for VR [31], and we expect that in the future, we will see players wear cameras to provide first-person viewpoints of a game. Since the average player height in the 2018-2019 season was over 200cm [38], which is outside the "acceptable" range of camera heights found here, we suspect a drop-off in terms of user experience. More work is needed to help understand how people react to large gaps in height between viewer and streamer.

5.4 Limitations and Future Work

We would like to acknowledge the limitations of our work that can inform future research. First, our study merged related work in immersive body-worn telepresence within a virtual reality context. Therefore, our findings inform researchers and designers on understanding users' camera height preference in the context of a simulated virtual environment. However, while this work might inform future work on human-to-human immersive telepresence, we cannot verify the generalization of our findings to the real world. As such, an area of future inquiry would be to replicate this study in a real context, instead of a VE, to see if our results withstand.

Additionally, within the SAP community, distance estimation tasks are usually constructed through blind-walking, bean-bag tossing, or verbal judgment calls [14, 29, 32, 37]. We are more interested in general scenarios moving towards immersive telepresence, and here offer a more user-centric measurement. However, our study did not include a measure for user satisfaction. Rather, we used self-selected camera height as a proxy for their personal preference. In our future studies, we would like to incorporate both user-selected choice and measures of user satisfaction to better compare our results to the previous findings of Keskinen et al. [16]. Further, in our study, participants were asked to remain seated, as the disconnect between the visual and vestibular systems could have caused them to lose balance and fall. As such, we did not measure if a viewer's position (seated vs. standing) had an effect on participant response. In future work, we plan on designing a safe procedure to measure this potential outcome.

Lastly, we acknowledge that in our study and in previous studies ([16, 41]), the majority of participants were male. Though our findings suggest that gender has a statistical influence on camera height preference, we note that our sample of 93 participants had very little overlap between male and female sub-populations. As such, an implication for future work is to perform stratified sampling near the mean of participant height for men and women, so to ensure that there is not a drastic difference in sub-populations. In addition, we need to consider other individual differences not identified by our work. As immersive telepresence is projected to be used by a wide variety of people, it is important to understand how different populations will have different camera height preferences.

6 CONCLUSION

Previous research on setting camera height in 360° film and immersive VR tried to find "one-size-fits-all" values, but we have shown that there are multiple factors which influence this design decision. VR and telepresence designers should consider the social landscape of the intended scene. Additionally, we show that the natural physiological difference between men and women may warrant different heights choices. In a telepresence scenario, we expect that we can accommodate these differences by having a streamer wearing a camera on different parts of the body, without affecting quality of view in terms of height. Overall, in VR and in social telepresence applications, there seems to be a wide buffer of acceptable camera heights which can still accommodate a natural viewing experience. Therefore, we do not anticipate a loss of quality with a small deviation within this buffer.

ACKNOWLEDGMENTS

This work is supported in part by Army RDECOM Award W911QX13C0052 and NSF Awards IIS-1638060 and IIS-1917728. We also thank the anonymous reviewers for their insightful feedback.

REFERENCES

[1] Albara Alohali, Kai Kunze, and Robert Earle. 2016. Run with me: designing storytelling tools for runners. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM, 5–8.

[2] S Aseeri, K Paraiso, and V Interrante. 2019. Investigating the Influence of Virtual Human Entourage Elements on Distance Judgments in Virtual Architectural Interiors. *Front. Robot. AI 6: 44. doi: 10.3389/frobt* (2019).

[3] Uddipana Baishya and Carman Neustaedter. 2017. In Your Eyes: Anytime, Anywhere Video and Audio Streaming for Couples. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 84–97.

[4] Deborah C. Beidel, B. Christopher Frueh, Sandra M. Neer, Clint A. Bowers, Benjamin Trachik, Thomas W. Uhde, and Anouk Grubaugh. 2019. Trauma management therapy with virtual-reality augmented exposure therapy for combat-related PTSD: A randomized controlled trial. *Journal of Anxiety Disorders* 61 (2019), 64 – 74. https://doi.org/10.1016/j.janxdis.2017.08.005 Virtual reality applications for the anxiety disorders.

[5] Minghao Cai, Soh Masuko, and Jiro Tanaka. 2018. Shopping Together: A Remote Co-shopping System Utilizing Spatial Gesture Interaction. In *Human-Computer Interaction. Interaction Technologies*, Masaaki Kurosu (Ed.). Springer International Publishing, Cham, 219–232.

[6] Minghao Cai and Jiro Tanaka. 2017. Trip Together: A Remote Pair Sightseeing System Supporting Gestural Communication. In *Proceedings of the 5th International Conference on Human Agent Interaction* (Bielefeld, Germany) *(HAI Š17)*. Association for Computing Machinery, New York, NY, USA, 317Ű324. https://doi.org/10.1145/3125739.3125762

[7] Brittany A Duncan and Robin R Murphy. 2017. Effects of Speed, Cyclicity, and Dimensionality on Distancing, Time, and Preference in Human-Aerial Vehicle Interactions. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 7, 3 (2017), 13.

[8] Cheryl D Fryar, Qiuping Gu, Cynthia L Ogden, and Katherine M Flegal. 2016. Anthropometric reference data for children and adults; United States, 2011-2014. *Vital Health Stat* 3, 39 (2016). https://stacks.cdc.gov/view/cdc/12223

[9] Clarissa Ishak, Carman Neustaedter, Dan Hawkins, Jason Procyk, and Michael Massimi. 2016. Human proxies for remote university classroom attendance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 931–943.

[10] Shunichi Kasahara, Shohei Nagai, and Jun Rekimoto. 2017. Jackin head: Immersive visual telepresence system with omnidirectional wearable camera. *IEEE transactions on visualization and computer graphics* 23, 3 (2017), 1222–1234.

[11] Shunichi Kasahara and Jun Rekimoto. 2014. JackIn: integrating first-person view with out-of-body vision generation for human-human augmentation. In *Proceedings of the 5th Augmented Human International Conference*. ACM, 46.

[12] Shunichi Kasahara and Jun Rekimoto. 2015. Jackin head: An immersive human-human telepresence system. In *SIGGRAPH Asia 2015 Emerging Technologies*. ACM, 14.

[13] Tadakazu Kashiwabara, Hirotaka Osawa, Kazuhiko Shinozawa, and Michita Imai. 2012. TEROOS: a wearable avatar to enhance joint activities. In *Proceedings of the*

SIGCHI Conference on Human Factors in Computing Systems. ACM, 2001–2004.
[14] Jonathan W. Kelly, Lucia A. Cherep, and Zachary D. Siegel. 2017. Perceived Space in the HTC Vive. *ACM Trans. Appl. Percept.* 15, 1, Article 2 (July 2017), 16 pages. https://doi.org/10.1145/3106155
[15] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology* 3, 3 (1993), 203–220. https://doi.org/10.1207/s15327108ijap0303_3
[16] T. Keskinen, V. Mäkelä, P. Kallioniemi, J. Hakulinen, J. Karhu, K. Ronkainen, J. Mäkelä, and M. Turunen. 2019. The Effect of Camera Height, Actor Behavior, and Viewer Position on the User Experience of 360ř Videos. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 423–430. https://doi.org/10.1109/VR.2019.8797843
[17] Jangyoon Kim and Victoria Interrante. 2017. Dwarf or Giant: The Influence of Interpupillary Distance and Eye Height on Size Perception in Virtual Environments. In *Proceedings of the 27th International Conference on Artificial Reality and Telexistence and 22Nd Eurographics Symposium on Virtual Environments* (Adelaide, Australia) *(ICAT-EGVE '17)*. Eurographics Association, Goslar Germany, Germany, 153–160. http://dl.acm.org/citation.cfm?id=3298830.3298859
[18] Don Kimber, Patrick Proppe, Sven Kratz, Jim Vaughan, Bee Liew, Don Severns, and Weiqing Su. 2014. Polly: Telepresence from a GuideŠs Shoulder. In *European Conference on Computer Vision*. Springer, 509–523.
[19] Sven Kratz, Daniel Avrahami, Don Kimber, Jim Vaughan, Patrick Proppe, and Don Severns. 2015. Polly Wanna Show You: Examining Viewpoint-Conveyance Techniques for a Shoulder-Worn Telepresence System. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. ACM, 567–575.
[20] Sven Kratz, Don Kimber, Weiqing Su, Gwen Gordon, and Don Severns. 2014. Polly: Being there through the parrot and a guide. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*. ACM, 625–630.
[21] Eike Langbehn, Gerd Bruder, and Frank Steinicke. 2016. Scale matters! Analysis of dominant scale estimation in the presence of conflicting cues in multi-scale collaborative virtual environments. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 211–220.
[22] Chris Lavigne. 2016. *360 Video Production Tactics: What We've Learned So Far*. Retrieved March 3, 2018 from https://wistia.com/learn/production/360-video-shooting-techniques
[23] Joseph J LaViola Jr, Ernst Kruijff, Ryan P McMahan, Doug Bowman, and Ivan P Poupyrev. 2017. *3D user interfaces: theory and practice*. Addison-Wesley Professional.
[24] Markus Leyrer, Sally A. Linkenauger, Heinrich H. Bülthoff, Uwe Kloos, and Betty Mohler. 2011. The Influence of Eye Height and Avatars on Egocentric Distance Estimates in Immersive Virtual Environments. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization* (Toulouse, France) *(APGV '11)*. ACM, New York, NY, USA, 67–74. https://doi.org/10.1145/2077451.2077464
[25] Javier Marín-Morales, Juan Luis Higuera-Trujillo, Carla De-Juan-Ripoll, Carmen Llinares, Jaime Guixeres, Susana Iñarra, and Mariano Alcañiz. 2019. Navigation Comparison between a Real and a Virtual Museum: Time-dependent Differences using a Head Mounted Display. *Interacting with Computers* 31, 2 (07 2019), 208–220. https://doi.org/10.1093/iwc/iwz018 arXiv:http://oup.prod.sis.lan/iwc/article-pdf/31/2/208/29109172/iwz018.pdf
[26] Marvin Minsky. 1980. Telepresence. Omni Magazine. *New York, Jun* (1980).
[27] Kana Misawa and Jun Rekimoto. 2015. Chameleonmask: Embodied physical and social telepresence using human surrogates. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 401–411.
[28] Kana Misawa and Jun Rekimoto. 2015. Wearing another's personality: a human-surrogate system with a telepresence face. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*. ACM, 125–132.
[29] Betty J. Mohler, Sarah H. Creem-Regehr, and William B. Thompson. 2006. The Influence of Feedback on Egocentric Distance Judgments in Real and Virtual Environments. In *Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization* (Boston, Massachusetts, USA) *(APGV '06)*. ACM, New York, NY, USA, 9–14. https://doi.org/10.1145/1140491.1140493
[30] Betty J. Mohler, Sarah H. Creem-Regehr, William B. Thompson, and Heinrich H. Bülthoff. 2010. The Effect of Viewing a Self-avatar on Distance Judgments in an Hmd-based Virtual Environment. *Presence: Teleoper. Virtual Environ.* 19, 3 (June 2010), 230–242. https://doi.org/10.1162/pres.19.3.230
[31] NBA. 2019. *NBA VR and MR*. Retrieved May 7, 2019 from https://www.nba.com/xr
[32] Karla Paraiso and Victoria Interrante. 2017. Can Virtual Human Entourage Elements Facilitate Accurate Distance Judgments in VR?. In *Virtual Reality and Augmented Reality*. Springer International Publishing, 119–133.
[33] Kevin Pfeil, Pamela Wisniewski, and Joseph LaViola Jr. 2019. An Analysis of User Perception Regarding Body-Worn 360° Camera Placements and Heights. In *ACM Symposium on Applied Perception 2019* (Barcelona, Spain) *(SAP '19)*. ACM, New York, NY, USA, Article 13, 10 pages. https://doi.org/10.1145/3343036.3343120
[34] L. Phillips and V. Interrante. 2011. A little unreality in a realistic replica environment degrades distance estimation accuracy. In *2011 IEEE Virtual Reality Conference*. 235–236. https://doi.org/10.1109/VR.2011.5759485
[35] Emily Price. 2017. *How to Shoot Amazing Pictures with a 360-Degree Camera*. Retrieved March 3, 2018 from https://lifehacker.com/how-to-shoot-amazing-pictures-with-a-360-degree-camera-1795928246
[36] Eric Ragan, Curtis Wilkes, Doug A Bowman, and Tobias Hollerer. 2009. Simulation of augmented reality systems in purely virtual environments. In *2009 IEEE Virtual Reality Conference*. IEEE, 287–288.
[37] E. D. Ragan, C. Wilkes, Y. Cao, and D. A. Bowman. 2012. The effects of virtual character animation on spatial judgments. In *2012 IEEE Virtual Reality Workshops (VRW)*. 141–142. https://doi.org/10.1109/VR.2012.6180921
[38] RealGM. 2019. *NBA Players*. Retrieved May 7, 2019 from https://basketball.realgm.com/nba/players
[39] Rebekka S. Renner, Boris M. Velichkovsky, and Jens R. Helmert. 2013. The Perception of Egocentric Distances in Virtual Environments - A Review. *ACM Comput. Surv.* 46, 2, Article 23 (Dec. 2013), 40 pages. https://doi.org/10.1145/2543581.2543590
[40] Sylvia Rothe, Boris Kegeles, Mathias Allary, and Heinrich Hußmann. 2018. The impact of camera height in cinematic virtual reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*. ACM, 124.
[41] Sylvia Rothe, Boris Kegeles, and Heinrich Hussmann. 2019. Camera Heights in Cinematic Virtual Reality: How Viewers Perceive Mismatches Between Camera and Eye Height. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video* (Salford (Manchester), United Kingdom) *(TVX '19)*. ACM, New York, NY, USA, 25–34. https://doi.org/10.1145/3317697.3323362
[42] Monique M Samson, Alan Crowe, PL De Vreede, Jos AG Dessens, Sijmen A Duursma, and Harald JJ Verhaar. 2001. Differences in gait parameters at a preferred walking speed in healthy subjects due to age, height and body weight. *Aging clinical and experimental research* 13, 1 (2001), 16–21.
[43] MHD Saraiji, Tomoya Sasaki, Reo Matsumura, Kouta Minamizawa, and Masahiko Inami. 2018. Fusion: full body surrogacy for collaborative communication. In *ACM SIGGRAPH 2018 Emerging Technologies*. ACM, 7.
[44] Richard M Satava. 1993. Virtual reality surgical simulator. *Surgical endoscopy* 7, 3 (1993), 203–205.
[45] Nova Shade. 2018. . Retrieved May 15, 2019 from https://assetstore.unity.com/packages/3d/environments/showroom-environment-73740
[46] Mel Slater, Vasilis Linakis, Martin Usoh, and Rob Kooper. 1996. Immersion, Presence and Performance in Virtual Environments: An Experiment with Tri-Dimensional Chess. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (Hong Kong) *(VRST Š96)*. Association for Computing Machinery, New York, NY, USA, 163Ű172. https://doi.org/10.1145/3304181.3304216
[47] Jamie Stark. 2017. *12 tips for shooting and editing 360-degree video for journalism*. Retrieved March 3, 2018 from https://medium.com/@StanfordJournalism/12-tips-for-shooting-and-editing-360-degree-video-for-journalism-4972cf50b77d
[48] Anthony Tang, Omid Fakourfar, Carman Neustaedter, and Scott Bateman. 2017. Collaboration with 360° Videochat: Challenges and Opportunities. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (Edinburgh, United Kingdom) *(DIS '17)*. ACM, New York, NY, USA, 1327–1339. https://doi.org/10.1145/3064663.3064707
[49] William B. Thompson, Peter Willemsen, Amy A. Gooch, Sarah H. Creem-Regehr, Jack M. Loomis, and Andrew C. Beall. 2004. Does the Quality of the Computer Graphics Matter when Judging Distances in Visually Immersive Environments? *Presence: Teleoperators and Virtual Environments* 13, 5 (2004), 560–571. https://doi.org/10.1162/1054746042545292
[50] Hiroaki Tobita. 2017. Gutsy-Avatar: Computational Assimilation for Advanced Communication and Collaboration. In *Robotic Computing (IRC), IEEE International Conference on*. IEEE, 8–13.
[51] Yuichi Tsumaki, Fumiaki Ono, and Taisuke Tsukuda. 2012. The 20-DOF miniature humanoid MH-2: A wearable communication system. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 3930–3935.
[52] Bob G Witmer and Michael J Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence* 7, 3 (1998), 225–240.
[53] I. Yavrucuk, E. Kubali, and O. Tarimci. 2011. A low cost flight simulator using virtual reality tools. *IEEE Aerospace and Electronic Systems Magazine* 26, 4 (April 2011), 10–14. https://doi.org/10.1109/MAES.2011.5763338