The Visual Computer
https://doi.org/10.1007/s00371-025-03934-5

RESEARCH

Evaluating behavioral realism in AR and VR: a comparison of single-point IK and full-body motion capture virtual humans

Elhassan Makled[1] · Christoph Gerhardt[1] · Tobias Schwandt[1] · Florian Weidner[2] · Wolfgang Broll[1]

Accepted: 21 April 2025

Abstract

Behavioral realism plays a crucial role in virtual and augmented reality (VR/AR). Various avatar animation techniques, ranging from full-body motion capture to single-point inverse kinematics (IK), offer different levels of realism. While the animation of a user's own avatar influences embodiment, the perceived realism of others' avatars is equally important for immersion. This study ($N = 53$) examines how users in smartphone AR, head-mounted display (HMD) AR, and VR perceive the behavioral realism of avatars animated with single-point IK compared to those driven by full-body motion capture. In addition, we explore whether the congruence between visual fidelity of an avatar and tracking accuracy affects perception. Our findings indicate that full-body motion capture produces significantly higher perceived realism than single-point IK, but the type of device does not have measurable impact. Furthermore, while congruence between visual realism and tracking fidelity was expected to play a role, our results suggest that its influence is limited. Despite lower realism than motion capture, modern IK techniques are still perceived positively, highlighting their viability for multi-user AR and VR applications.

Keywords Virtual reality · Augmented reality · Avatars · Behavioral realism

1 Introduction

AR and VR technologies have garnered considerable interest in recent years. Social applications within these domains aim to connect people and introduce novel forms of interaction. A fundamental element of such applications is user representation. The 3D models used for avatars can adopt diverse appearances, ranging from highly abstract to photorealistic. These variations extend to rendering styles (e.g., abstract, cartoon-like, or realistic) as well as the inclusion of body parts, which may be limited to hands and head, or a full-body depiction [48]. Studies have shown that in most cases realistic full-body representations yield superior results in terms of presence [39, 55], task performance [15, 25], and communication behavior [14, 47]. When designed in a realistic and anthropomorphic style, these representations typically exhibit human-like movements. However, achieving a high level of behavioral realism in full-body animations demands considerable effort and meticulous implementation.

Behavioral realism refers to the degree to which human avatars exhibit naturalistic behaviors in real time, aligning with human expectations in physical environments [7, 42]. A straightforward approach to achieving high behavioral realism involves tracking each joint and bone in the human body using motion capture systems such as OptiTrack[1] or ART.[2] Although these systems offer comprehensive body tracking, they require specialized and often expensive hardware, which is impractical for many real-world applications. To reduce reliance on such extensive setups, inverse kinematics (IK) is employed to approximate the positions of untracked joints using a limited set of data points, such as the hands, feet,

✉ Elhassan Makled
elhassan.makled@tu-ilmenau.de

Christoph Gerhardt
christoph.gerhardt@tu-ilmenau.de

Tobias Schwandt
tobias.schwandt@tu-ilmenau.de

Florian Weidner
f.weidner@lancaster.ac.uk

Wolfgang Broll
wolfgang.broll@tu-ilmenau.de

1 Ilmenau University of Technology, Ilmenau, Germany

2 Lancaster University, Lancaster, UK

[1] OptiTrack, NaturalPoint Inc., https://optitrack.com, last accessed: 22.03.2025.

[2] ART, ART, https://ar-tracking.com, last accessed: 22.03.2025.

hips, and head [8]). IK estimates these missing joint positions by analyzing key aspects of human anatomy, including joint degrees of freedom, range of motion, and bone length [2] thereby calculating the most plausible joint placement.

In the most extreme cases, full-body animation is generated from a single data point, a method known as single-point IK. Although this approach is less relevant for HMD-based applications—since these typically include hand tracking or tracked controllers that enable at least three-point tracking—it becomes highly significant in smartphone-based AR. In such scenarios, the only available tracking point is usually the smartphone itself. Previous research by Ahuja et al. [1] and Makled et al. [23] investigated the use of a smartphone as a single tracking point to drive IK-based animations. Their research indicated that utilizing a single data point is a practical strategy and that self-avatars animated through this method can create a significant feeling of embodiment. However, both studies highlight that tracking accuracy declines as the distance from the smartphone (the tracking point) increases.

Building upon these findings, this study offers results from an experiment aimed at assessing whether the perception of *others* demonstrates an adequately high degree of behavioral realism. Gaining insight into how users perceive the behavioral realism of IK-animated avatars is essential for designing VR and AR applications that foster user engagement and enhance immersion. For instance, insufficient behavioral realism can reduce social presence [20] and hinder effective non-verbal interpersonal communication. [19].

In our study, participants used either a VR HMD, a pass-through AR HMD, or an AR-capable smartphone to explore a museum artifact while a virtual human was present. The virtual human stood nearby, engaging with the virtual space and the artifact. The setup of the study is illustrated in Fig. 2. Participants were not explicitly instructed to observe or evaluate the movements of the virtual human. Instead, the goal was to examine perception in scenarios where others are simply bystanders, similar to a real museum setting. The virtual human was animated using two types of pre-recorded tracking data: one based on full-body motion capture and the other generated through a single-point IK algorithm with the smartphone as input. Participants experienced both conditions and subsequently completed a questionnaire assessing behavioral realism. Furthermore, since previous research suggests that behavioral realism interacts with visual realism [4, 41], we examine how these factors influence each other. In particular, we investigate whether a low-realism visual representation performs better when paired with less precise tracking data, such as single-point IK. To explore this, we conducted a mixed-design study comparing single-point IK and full-body tracking across HMD-based VR, pass-through HMD AR and smartphone-based AR, using both abstract and realistic avatar representations.

Our results show that current IK animations are a competitive solution when it comes to animating virtual humans in VR and smartphone AR. Although we found a significant difference between both, with single-point IK being attributed to less behavioral realism than motion capture, the difference and effect are small ($\Delta = 0.24$ on a scale of 1–6, $\eta_p^2 = 0.042$). This has practical implications for smartphone AR and HMD VR, as it suggests that social HMD-based VR applications do not require complex tracking solutions for full-body animations to achieve reasonable levels of behavioral realism of others.

Similarly, it shows that single-point IK, relying only on the smartphone's tracking data as information, enables smartphone AR users to partake in social AR applications, experiencing others with behavioral realism comparable to systems that use full-body tracking data.

The remainder of this paper is structured as follows: In Sect. 2, we review related work on virtual humans, behavioral realism, and tracking, including inverse kinematics (Inverse Kinematics (IK)). We then describe our study design and procedure in Sect. 3, followed by the presentation of our results in Sect. 4. Subsequently, in Sect. 5, we provide context for our findings, discuss our limitations, and, finally, in Sect. 6, we draw conclusions from our work.

2 Related work

This section examines prior research on virtual humans in VR and AR applications, focusing on various tracking methods employed to embody users within these systems.

2.1 Virtual and augmented realities

Virtual reality (VR) and augmented reality (AR) are immersive technologies that enable users to interact, experience telepresence, and become immersed in digital environments [30]. In recent years, virtual reality (VR) has been widely adopted across industries such as entertainment, healthcare, education, and tourism [10, 24, 27]. It has also facilitated social interactions in virtual spaces [51], made possible through head-mounted displays (head-mounted displays (HMDs)) that fully immerse users in digital environments.

Similarly, augmented reality (AR) blends digital content with the real world, allowing users to interact with virtual objects while remaining aware of their physical surroundings. Unlike VR, which requires dedicated headsets, AR can run on various devices, with smartphones being the most common platform. Popular applications such as Instagram[3]

[3] Instagram, Meta, https://instagram.com, last accessed: 25.03.2025.

and Pokémon Go[4] exemplify its widespread use. However, AR also supports HMDs like Magic Leap and Meta Quest 3, which utilize either *pass-through* or *see-through* technology.

Pass-through AR relies on cameras to capture and display the user's surroundings inside the headset, while *see-through* AR uses semi-transparent displays to overlay digital graphics onto the real world. A key limitation of current see-through devices is their restricted field of view. For instance, the Magic Leap 2 offers a maximum field of view of 70 degrees,[5] whereas Meta's Quest 3 provides a broader 110-degree field of view[6] for pass-through experiences.

In our study, we utilize pass-through AR due to its wider field of view, which allows users to observe augmented digital objects more effectively.

2.2 Virtual humans

In numerous VR applications, users navigate digital environments by embodying virtual humans as their avatars, a concept explored by Weidner et al. [48]. Among other use cases, embodying a virtual human is a common practice in both single-user and multi-user VR applications. These virtual humans serve as interfaces, facilitating communication with other users and enabling interactions within the virtual environment [9, 13, 34, 35].

When examining user perception of virtual humans in virtual environments, researchers typically categorize them into two contexts: computer controlled or user controlled. Computer-controlled virtual humans are often referred to as agents or non-player characters (NPCs) [6, 18, 47], whereas user-controlled virtual humans are commonly known as avatars [43, 48]. Avatars are deeply connected to users' physical selves [11, 33, 36, 37, 52], fostering a sense of presence and embodiment within virtual environments [46]. By serving as a bridge between users and the virtual world, avatars facilitate interaction and engagement [33], enhancing authenticity and realism. This, in turn, improves cognitive capabilities [40] and haptic performance [16, 28]. Makled et al. [23] demonstrated that avatars animated using single-point IK can evoke similar levels of embodiment (self-perception) compared to those animated through full-body motion capture. In our study, we examine whether a virtual human controlled by single-point IK can achieve a comparable—or even higher—degree of behavioral realism than one driven by full-body tracking data. Although users are actually viewing agents, they are led to believe that another human controls the virtual human, creating the impression of an avatar.

Wang et al. [47] investigated different agent representations in AR using the Microsoft HoloLens 2. They presented users with three variations: a full-size human, a miniature human, and a non-human form (a cylinder), alongside a voice-only condition. Their results indicated that participants responded more positively to human-like representations. Similarly, research by Wagner et al. [44] and Wainer et al. [45] found that users reported greater enjoyment and amusement when interacting with embodied agents that closely resembled humans. However, the perception of virtual humans is influenced not only by their visual appearance but also by their behavior.

The following section will review previous research on behavioral realism in avatars and its impact on user perception.

2.3 Behavioral realism

Behavioral realism refers to the degree to which human avatars can exhibit natural behaviors in real-time, accurately reflecting the expected actions of humans in the physical world [19]. Behavioral realism is vital in shaping user experiences and interpersonal interactions within virtual environments. Accurate representation of avatar movements, facial expressions, body language, and gaze behavior improves engagement, immersion, trust, and rapport among users [3]. Creating virtual interactions that emphasize behavioral realism can improve user satisfaction and enhance the overall effectiveness of virtual communication platforms [19].

Herrera et al. [19] explored this phenomenon further by comparing avatars with different levels of behavioral realism. Their findings revealed that participants experienced improved interpersonal results, as well as increased self-presence and interpersonal attraction, when they had greater control over avatar movements. This indicates that the accuracy of avatar behavior plays a crucial role in shaping users' perceptions and interactions within virtual environments.

Groom et al. [18] investigated behavioral realism through the frameworks of Uncanny Valley Theory, Maximization Theory, and Consistency Theory. Consistency theory suggests that virtual humans with behaviors that remain consistently realistic at a given level are preferred over those with inconsistencies. Their study focused on how different levels of behavioral realism, such as the inclusion or omission of lip sync and body movement, affect human interactions. However, they did not examine the impact of consistency between the visual representation of a virtual human and its behavior (animations).

Expanding the discussion to AR, Jun et al. [20] examined the influence of behavioral realism and anthropomorphic realism (the visual fidelity of virtual humans) on social pres-

[4] Pokémon Go, Niantic, https://pokemongolive.com/, last accessed: 25.03.2025.

[5] Magic Leap 2, Magic Leap, https://www.magicleap.com/optics, last accessed: 25.03.2025.

[6] Quest 3, Meta, https://www.meta.com/quest/quest-3/, last accessed: 25.03.2025.

ence. Their findings revealed that participants experienced a stronger sense of presence with human-shaped objects that exhibited both higher anthropomorphic realism and greater behavioral realism. This underscores the complex relationship between visual and behavioral fidelity in shaping user perceptions and interactions within virtual environments.

In the domain of immersive virtual environments, research by Kyrlitsias and Michael-Grigoriou [21] and Yu et al. [53] highlighted the pivotal role of behavioral realism in eliciting natural responses and fostering social interactions. Both studies emphasized that accurately depicting avatar behavior significantly improves social presence and engagement in virtual environments.

Achieving a high level of behavioral realism requires accurately mapping user movements to their avatars, which is facilitated by various tracking technologies. The following section will review previous research on tracking methods that enable users to control avatar movements and maintain a sense of agency.

2.4 Tracking & inverse kinematics

As previously discussed in Sect. 2.2, Groom et al. [18], along with Wang et al. [47] and Bevacqua et al. [6], highlight the importance of ensuring that an avatar's behavior aligns with its visual appearance. Building on this, an avatar with a realistic rendering style should exhibit equally realistic behavior or tracking. To achieve natural and believable movement, a user's actions must be accurately mapped to their avatar in real time. This can be accomplished by tracking the user's body using external sensors such as cameras [50] or depth sensors [54], or by leveraging built-in tracking capabilities of devices like smartphones and HMDs [23, 31, 49]. Most commercially available HMDs, such as the Apple Vision Pro[7] and the Meta Quest 3, typically track only three body points: the user's head and hands. To animate the rest of the body, applications such as VRChat[8] and Meta Horizon Worlds[9] rely on inverse kinematics (IK) to estimate untracked movements.

Some HMDs, like the HTC Vive,[10] allow users to enhance tracking fidelity by adding external trackers. Eubanks et al. [11] explored how different combinations of tracked body points—such as the head, hands, hips, and feet—affect user embodiment. Their findings suggest that tracking additional points significantly enhances the sense of embodiment, particularly when feet tracking is included.

While multi-point tracking improves realism, smartphone-based AR lacks external sensors, making full-body tracking infeasible. To address this, researchers have developed single-point IK approaches, where only a single tracking point (e.g., a smartphone) drives the animation of the entire avatar. Makled et al. [23], Ahuja et al. [1], and Murugan et al. [32] have demonstrated the viability of this method to animate avatars in collaborative virtual environments (CVE).

Different implementations of single-point IK vary in their approach. ARIKA by Makled et al. [23] and the method by Murugan et al. [32] utilize Unity's built-in IK solver for upper-body animation, supplementing it with pre-recorded walking and idle animations for the lower body. In contrast, Pose-on-the-Go by Ahuja et al. [1] incorporates additional sensor data from the iPhone XR.[11] Using the depth sensor on the phone, the system estimates chest orientation to improve avatar alignment. The IK itself is driven by the phone's inertial measurement unit (IMU), which provides position and orientation data to animate the full body.

Makled et al. [23] further examined the impact of single-point IK avatars on user embodiment, finding that while motion-captured avatars provide a stronger sense of agency and self-location, there is no significant difference in body ownership compared to single-point IK avatars. These results suggest that single-point IK can effectively support self-embodiment in AR.

Previous work has primarily focused on how IK affects the users controlling the avatars, often overlooking the perspective of external observers. However, in multi-user AR and VR scenarios, it is equally important to understand how others perceive IK-animated avatars. This study aims to address this gap by investigating the behavioral realism of single-point IK avatars in both smartphone-based AR and HMD-based VR, considering not only user embodiment but also how these avatars are perceived by bystanders.

To investigate the perceived behavioral realism of IK-driven virtual humans, we draw upon Consistency Theory [18] (cf. Sect. 2.3) to ensure alignment between the visual representation of a virtual human and its tracking method. Specifically, for our ground truth condition (motion-captured tracking), we employ a realistic human avatar, whereas for the IK-driven condition, we use an abstract robotic avatar to reflect the computational nature of IK-generated animations (cf. Fig. 3).

We aim to examine whether an interaction exists between single-point IK and the visual realism of the virtual human, influencing users' perception of its behavioral realism.

[7] Vision Pro, Apple, https://www.apple.com/apple-vision-pro/, last accessed: 23.03.2025.

[8] VRChat, VRChat, https://hello.vrchat.com

[9] Horizon Worlds, Meta, https://www.meta.com/en-gb/experiences/2532035600194083/, last accessed: 23.03.2025.

[10] Vive, HTC, https://vive.com, last accessed: 23.03.2025.

[11] iPhone XR, Apple, https://support.apple.com/111868, last accessed: 23.03.2025.

3 Study design

Building on the related work, we propose the following hypotheses:

H1: Behavioral realism will be perceived as highest when motion tracking is paired with a realistic avatar and when single-point IK is paired with an abstract avatar.
H2: Motion tracking-based animation will result in higher perceived behavioral realism compared to single-point IK.

Additionally, our objective is to explore whether the device type (VR HMD, AR via HMD, or AR via smartphone) influences perceived behavioral realism, potentially due to variations in field-of-view and interaction context.

The following sections detail our study design and the development of the IK system. The primary objective of our study is to assess how users perceive other virtual humans (rather than their own avatars) when those virtual humans are animated using single-point IK in VR and AR environments.

To examine this, we analyze the effects of different factors on the perceived behavioral realism of virtual humans in VR and AR. We employ a 2 × 2 × 3 experimental design, incorporating two within-subject variables—avatar realism (realistic vs. abstract) and animation source (motion tracking vs. single-point IK)—and one between-subject variable-device type (VR HMD, smartphone AR, HMD AR).

Although we considered making the device type a variable within the subject, we ultimately opted against it. Doing so would have significantly increased the duration of the experiment, required a complex counterbalancing procedure, and risked that participants repeatedly observed the same movement patterns, which could bias the scores and reduce their ability to differentiate between conditions. Thus, the investigated variables were:

- **Within-subject variables:**
  - *Visual realism:* The rendering style of the virtual human: {abstract; realistic}.
  - *Animation method:* The technique used to animate the virtual human: {single-point IK system; motion capture}.
- *Between-subject variable:*
  - **Device type:** The device used to visualize the virtual humans {VR; HMD AR; smartphone AR}.

As the independent variable, we measure perceived behavioral realism (cf. Sect. 3.4).

3.1 Apparatus

For the study, we developed a Unity-based application that enables participants to observe a virtual artifact alongside a virtual human animated by pre-recorded motion data. VR participants used a Meta Quest 2.[12] HMD AR participants used a Meta Quest 3 with pass-through functionality. Smartphone AR participants used a Google Pixel 6 Pro.[13]

During the study, AR participants (both smartphone and HMD users) viewed the virtual human and artifact through their respective devices, seeing them positioned on a pedestal within their real-world surroundings (cf. Figs. 1 and 3a). For VR participants, the entire scene, including the virtual human, artifact, pedestal, and even the surrounding room, was fully virtual (cf. Figs. 1 and 3b). To maintain consistency between conditions, we created a digital twin of the physical room setup used in the AR study, replicating it as the virtual environment for VR participants. Different environments are illustrated in Fig. 2.

3.1.1 Tracking

To achieve the core objective of our study—the evaluation of single-point IK—we utilized the IK system developed by Makled et al. [23]. This system was chosen due to its public accessibility and its status as the most recent single-point IK system for smartphone-based AR.

In the *single-point IK condition*, the full-body animation data are derived exclusively from tracking data provided by an Android smartphone and an IK algorithm. External sensors, such as an HMD or controllers, were not used. Although this approach enables full-body animation for users with only a smartphone (for example, in outdoor AR scenarios), it also introduces tracking inaccuracies, as most of movement is estimated by the IK algorithm. The average Euclidean distance error across all measurements is 22 cm (SD = 13 cm, min = 19.3 cm, max = 26.7 cm). The smallest error occurs at the right hand (which holds the smartphone), whereas the largest error is observed at the left foot. As a ground truth (and the second condition for the within-subject variable *animation method*), we used motion-captured animations recorded via an *OptiTrack* system.

3.1.2 Animations

Both the IK and motion capture animations were recorded simultaneously. A non-professional actor, equipped with a motion-tracking suit and holding a smartphone, was

[12] Quest 2, Meta, https://www.meta.com/quest/products/quest-2/, last accessed: 22.03.2025.

[13] Pixel 6 Pro, Google, https://store.google.com/de/category/phones, last accessed: 22.03.2025.

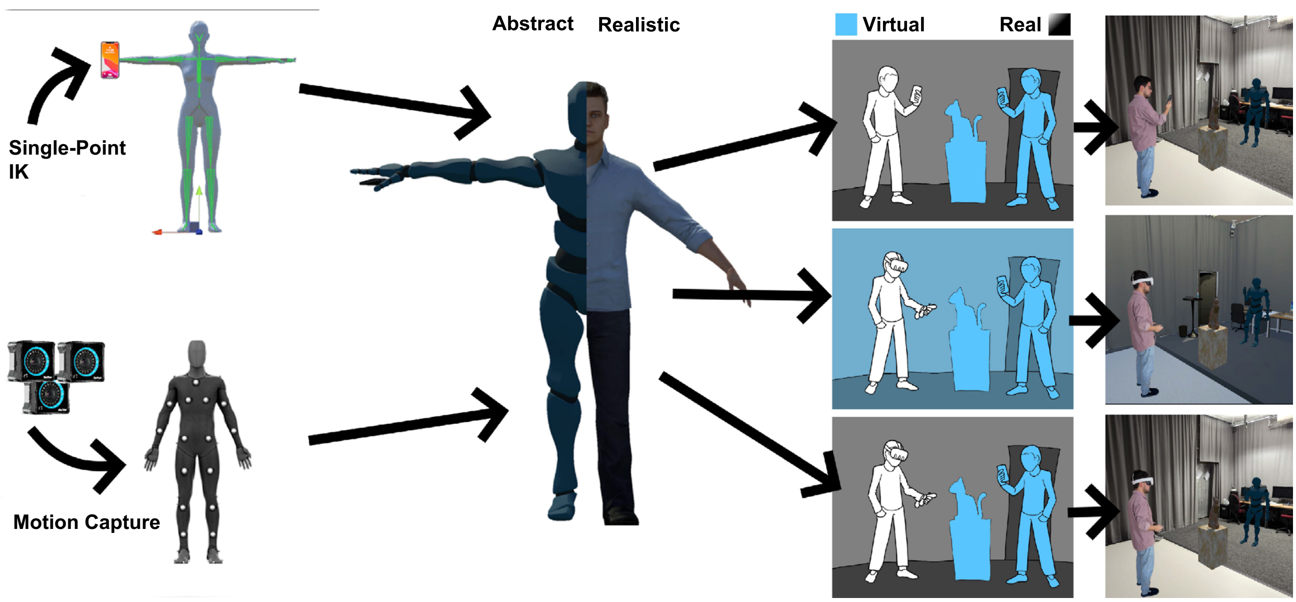


Fig. 1 Our study investigated single-point IK and motion capture as animation methods for both realistic and abstract virtual humans. Participants viewed these animations using a VR HMD, AR HMD, or smartphone AR, either in a physical laboratory environment (AR) or its virtual counterpart (VR)

instructed to perform predefined movements for the animations.

The animations consisted of the following movements: First, the virtual human remains stationary for approximately five seconds. Then, it moves sideways, taking three steps to the left and three to the right, followed by two steps forward and two backward. These movements simulate exploration, as if the virtual human is examining the artifact from different angles, with each step lasting around 2.5 s. Next, the user moves their smartphone in a circular motion, mimicking the action of scanning the environment or closely inspecting an object through the device. This animation sequence captures two key behaviors:

- Wide-range arm movements typical of smartphone AR users when analyzing an artifact.
- Body movement around the artifact, simulating spatial interaction.

To prevent repetition bias, the order of these movements was randomized using a Latin Square.

3.1.3 Avatars

As discussed in Sect. 2, the visual representation of a virtual human and the coherence between its appearance and animation can influence the perception of the observer, as suggested by *Consistency Theory*. To examine this effect, we introduce two distinct virtual human representations: abstract and realistic, as shown in Fig. 3. For the abstract virtual human, we used the Y-Bot model from Mixamo.[14] For the realistic virtual human, we selected the Male_Adult_8 model from Microsoft's Rocketbox [17]. These avatars were chosen for their accessibility to the research community and their previous use in studies on avatars and IK [5, 17, 26, 32, 48].

3.2 Task

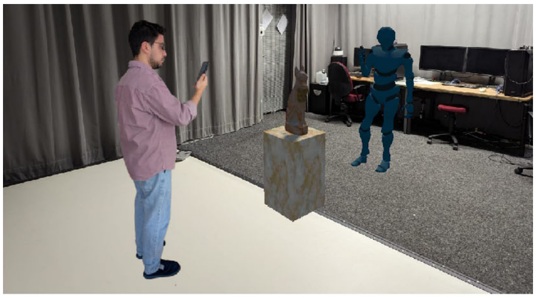
Participants were instructed to engage in an artifact inspection task regardless of rendering style, tracking method, or visualization condition.

At the beginning of each condition, they saw an artifact placed on a pedestal before them and could inspect it freely with six degrees of freedom. Care was taken to ensure that, upon starting a condition, participants saw the virtual human directly in front of them in both AR and VR (see Figs. 1 and 2).

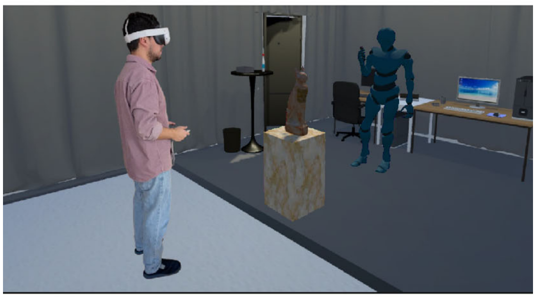
The virtual human was always displayed holding a smartphone, appearing to observe the space and artifact through the same AR application as the participant. Meanwhile, a second virtual human stood opposite the participant, walking and looking around, following a predefined animation sequence. The duration of the task was one minute and no additional instructions were provided.

We structured the study around an artifact inspection task to provide context for virtual human movements. Although a simpler setup, where participants were only asked to observe

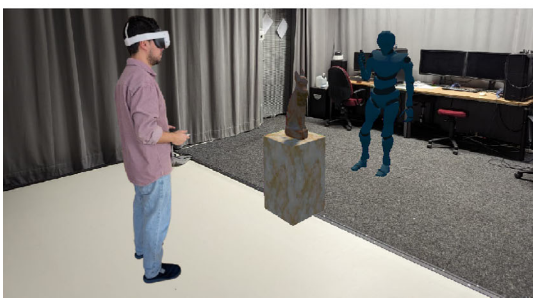
[14] Mixamo, Adobe, https://mixamo.com, last accessed: 23.03.2025.



(a) Smartphone AR



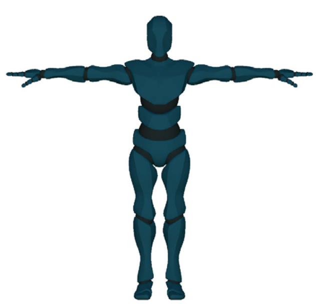
(b) VR



(c) HMD AR

Fig. 2 AR experiment setup (**a**) features a real-world room with an augmented pedestal and artifact, while the VR experiment setup (**b**) replicates this environment digitally to ensure consistency across conditions. HMD AR user setup (**c**), where the participant sees the virtual elements overlaid onto their physical surroundings

(a) Realistic



(b) Abstract (Y-Bot)

Fig. 3 Different virtual human representations used for visualizing the IK avatars included: (realistic avatar) Male_Adult_8, a detailed human-like model and (abstract avatar) a simplified, robotic representation designed to align with the computer-simulated nature of IK animations

the virtual human, was possible, we opted for a more immersive scenario.

This approach ensures that participants do not focus solely on critically assessing virtual human motion. Instead, the virtual human's behavioral realism contributes to their overall experience, reflecting how such avatars might function in real-world scenarios. Consequently, the results can be interpreted within the context of neutral to positive values, where a more realistic virtual human enhances the experience, whereas an unrealistic one may be perceived as distracting or disruptive.

3.3 Procedure

After providing informed consent, the participants completed a demographic questionnaire collecting data on age and gender identity. The participants were then briefed on the setup and hardware of the experiment. The experimenter explained the task, instructing the participants to imagine themselves in a museum, where they would be observing a virtual artifact alongside another person, a virtual human, who was also involved in the observation.

Participants were assigned to one of three device groups (VR HMD, HMD AR, or smartphone AR) and familiarized themselves with the respective hardware before starting the experiment. The study followed a within-subjects design, where each participant experienced four conditions, randomized using a complete Latin Square to counterbalance order effects. Thus, each participant experienced the following four unique conditions:

– Abstract avatar + single-point IK animation
– Abstract avatar + motion capture animation
– Realistic avatar + single-point IK animation
– Realistic avatar + motion capture animation

During each condition, the virtual human was positioned in front of the participant, performing the pre-recorded animations while both observed the artifact. This setup allowed us to examine how variations in avatar realism and animation technique influenced perceived behavioral realism.

Participants completed a questionnaire after each condition, assessing their perception of the virtual human's behavioral realism. Once all four conditions were met, they were thanked for their participation. The entire experiment lasted approximately 20 min per participant, including briefing, familiarization, task execution, and completion of the questionnaire.

3.4 Measures

To evaluate behavioral realism, we focus on how participants perceive single-point IK avatars compared to high-fidelity

Table 1 List of questions presented to participants to assess the behavioral realism of the virtual human

No.	Question
AP1	The avatar was not controlled by someone else
AP2	The avatar's behavior was believable
AP3	The avatar was behaving like a real person
AP4	The avatar was controlled by somebody else
AP5	The idea that the avatar did not behave like a real person crossed my mind
AP6	I perceived the avatar as a simple computer program

motion capture. Our measurement approach considers subjective ratings of realism, consistency between movement and appearance, and the avatar's integration within the scene. This allows us to assess differences in perception across various conditions.

The questionnaire by Bevacqua et al. [6] was adapted to assess the behavioral realism of avatars in our study. It includes items evaluating the naturalness, believability, and coherence of avatar movements in relation to their appearance and context. This approach allows us to quantify how well participants perceive the single-point IK avatars compared to motion-captured ones. To keep the duration of the experiment short, we only ask questions about behavioral realism and not task-related questions. The final questionnaire included items that assessed participants' perceptions of avatar realism, the naturalness of their movements, and their overall behavioral realism. These measures aimed to capture how convincingly the virtual humans' behavior aligned with real-world expectations. Table 1 presents a list of questions designed to assess the behavior realism of avatars.

Similar to the original study, participants responded to these questions using a 6-point Likert scale (1 is "disagree strongly" and 6 is "agree strongly"). The values of each trial were averaged to generate a final score.

3.5 Sample

A total of 53 participants participated in our study (18 for smartphone AR, 18 for HMD VR, and 17 for HMD AR). They were recruited via email sent to students and staff members of our university and did not receive any compensation. Among them, 32 self-identified as male, 20 as female, and 1 as agender.

The average age of the participants was 30.11 years (SD = 9.70), ranging from 20 to 70 years. Participants rated their interest in VR at $M = 4.27$ (SD = 0.87, min = 1, max = 5) and their interest in AR at $M = 4.06$ (SD = 0.98, min = 1, max = 5), using a 5-point Likert scale (1 = "not interested at all," 5 = "really interested"). The study received approval from the Institutional Review Board (IRB).

4 Results: behavioral realism of virtual humans

To assess significant differences among our three independent variables, we performed a mixed three-way ANOVA to examine the effects of style, animation, and visualization on behavioral realism.

- Visual realism (abstract or realistic)
- Animation method (IK or motion-captured)
- Device (VR or AR),

In our analysis, we computed a single average score by summing the responses from all questionnaire items and dividing by the total number of questions, accounting for negatively framed items such as AP1, AP5, and AP6. These items were inverted so that a response of 1 corresponded to "strongly agree" and 6 to "strongly disagree." Outliers were identified and addressed using the z-score method, resulting in the detection of six extreme cases [12]. Figure 4 illustrates the mean values and standard deviations for each question, categorized by our independent variables. Data generally followed a normal distribution ($p > 0.05$ in Shapiro–Wilk's test [38]).

To ensure that the assumptions for ANOVA were met, we assessed homogeneity of variances using Levene's test [22] ($p > 0.05$) and sphericity using Mauchly's test [29] ($p > 0.05$), both of which confirmed that the assumptions were satisfied.

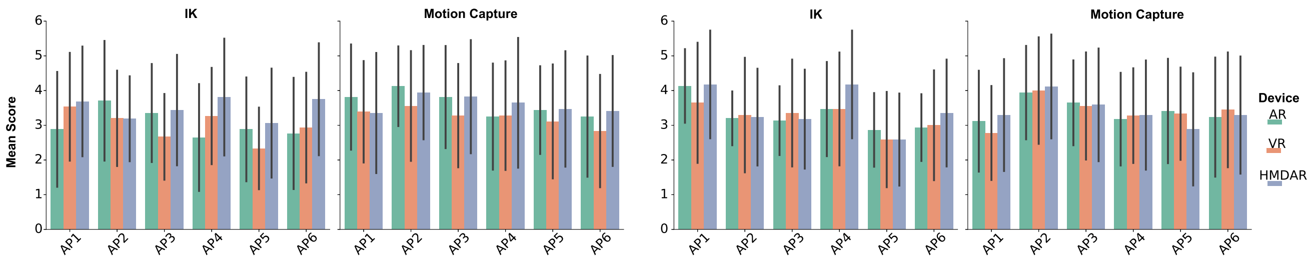
Table 2 presents the results of our statistical analysis. No significant differences were found in the two- and three-way interactions between device, visual realism, and animation method ($p > 0.05$). Furthermore, we did not observe a significant effect of device or visual realism (both $p > 0.05$). However, the animation method showed a significant main effect ($F(1,42) = 7.496$, $p = 0.009$, $\eta_p^2 = 0.05$). Specifically, the IK animation method scored $M_{IK} = 3.20$ (SD = 0.58), while the motion-captured animation method scored $M_{MC} = 3.47$ (SD = 0.63). Apart from this, no other significant differences were detected ($F(1,42) < 2.91$, $p > 0.07$, $\eta p^2 < 0.03$) (Fig. 5).

5 Discussion and limitations

Our findings offer key insights into the perception of single-point IK virtual humans in social AR and VR environments.

5.1 H1: effects of motion tracking and visual realism

Interestingly, our study found no significant differences in perceived behavioral realism based on visual realism (abstract vs. realistic). This suggests that the animation



(a) Abstract (b) Realistic

Fig. 4 Mean and standard deviation values of all behavioral realism questions Table 1 for different devices as well as IK and motion-captured animation types for - abstract avatar - realistic avatar (**b**) (min = 1, max = 6. The values are inverted for AP1, AP5, and AP6 as discussed in Sect. 4

Table 2 Three-way mixed ANOVA results for the three conditions concerning behavioral realism

	F	p	η_p^2
Device	$F(2, 42) = 1.592$	0.215	0.0320
VisualRealism	$F(1, 42) = 0.099$	0.755	0.0003
AnimationMethod *	**F(1, 42) = 7.496**	**0.009**	**0.0500**
Device × VisualRealism	$F(2, 42) = 2.906$	0.066	0.0180
Device × AnimationMethod	$F(2, 42) = 0.184$	0.833	0.0030
VisualRealism × AnimationMethod	$F(1, 42) = 0.436$	0.513	0.0010
Device:VisualRealism × AnimationMethod	$F(2, 42) = 0.491$	0.616	0.0030

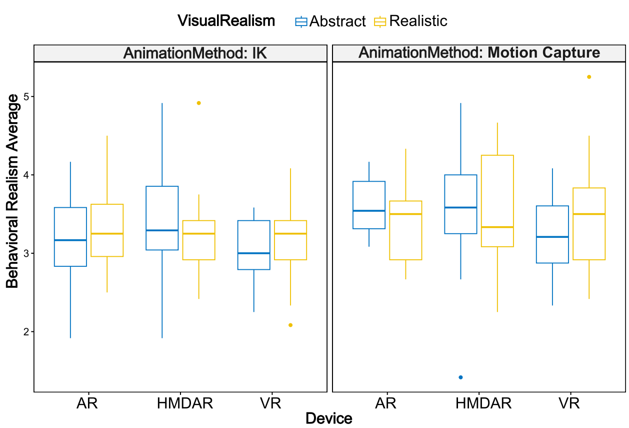


Fig. 5 Box plots of behavioral realism average by visual realism, colored by visual realism and faceted by animation method. We detected a significant main effect of animation method: Regardless of device and visual realism, motion capture led to higher behavioral realism. There were no other main effects or two- or three-way interaction effects

method plays a more critical role in user perception than the avatar's appearance or environmental fidelity.

Our findings reinforce that animation quality has a greater impact on perceived behavioral realism than visual realism. This aligns with previous research [18, 19], but challenges consistency theory [18], which suggests that visual and behavioral realism should be aligned. Several factors may explain this: (1) Users may have expected humanoid avatars-regardless of abstraction-to behave in a human-like manner. (2) Consistency theory also considers facial expressions and eye movements, which were absent in the realistic avatar, potentially reducing its perceived realism compared to the abstract one. This may explain why the abstract avatar sometimes scored higher when using motion capture.

5.2 H2: impact of motion tracking method

Our study compares the behavioral realism of virtual humans animated using full-body motion capture versus single-point IK. The results indicate that full-body motion capture generally leads to higher perceived behavioral realism.

However, our evaluation also found that single-point IK animations achieved a moderate level of behavioral realism (M =3.20, SD = 0.58). Although not as effective as motion capture, our findings suggest that modern IK solutions do not significantly degrade users' perception of behavioral realism in virtual humans. This underscores the viability of IK for multi-user AR and VR applications, particularly given the tracking limitations of smartphone AR devices.

5.3 Exploratory analysis: impact of device

The type of device (smartphone AR, HMD pass-through AR or HMD) did not significantly influence users' perceptions of avatars' behavioral realism. This suggests that there could be an interaction between device and visual realism. This is noteworthy given the inherent differences in field of view, tracking quality, and environmental immersion between smartphone AR, AR HMDs and VR HMDs.

While HMD users experience a wider virtual field of view, smartphone AR users retain constant awareness of the real world, reinforcing the absence of a physically present user. Despite these contrasts, the type of device did not affect the perceived behavioral realism within the shared space.

Although IK-based animations received moderately positive ratings, improvements are needed to better capture natural human movement. Smartphone AR users rely solely on the hand holding the phone to drive the IK system, with idle and walking animations compensating for missing body movements [23]. This omits key elements such as head movement, non-dominant hand gestures, and subtle posture shifts, factors accurately tracked by motion capture. Enhancing realism through sensor fusion, such as integrating smartwatches for hand tracking and front-facing cameras for facial tracking, could further bridge the gap between IK and full-body motion capture.

5.4 Limitations

This study has several limitations. First, the evaluation focused solely on avatar movement, without interaction or responsiveness to participants. Despite this, behavioral realism ratings were neutral to positive. However, excluding non-verbal cues and social interaction may have overlooked key influences on perceived realism.

Second, visual authenticity factors, such as the absence of environmental interactions (e.g., avatars casting shadows), were particularly limiting for AR users. Additionally, we did not track participants' gaze, leaving unknown how much attention was given to the avatar versus the artifact or surroundings. Finally, the sample size of 54 participants may limit the generalizability of our findings.

6 Conclusion and future work

Our study examined the impact of single-point inverse kinematics (IK) on the perceived behavioral realism of virtual humans in smartphone AR and HMD-based VR. A total of 54 participants observed realistic and abstract avatars animated using either motion tracking or IK. We explored how the congruence between visual realism and animation type influenced perception.

In summary, our results show that tracking type significantly affects behavioral realism, with motion-captured animations rated higher than IK animations. However, we did not find a significant impact of congruence or device type.

In conclusion, our findings show that full-body motion capture leads to higher perceived behavioral realism than single-point IK. However, modern single-point IK still achieves positive and competitive results. In addition, we found no significant interaction between avatar style, visualization device, and animation type to influence behavioral realism.

Future research should explore real-time multi-user interactions to assess the impact of behavioral realism in AR and VR. Additionally, future work should develop multiplayer scenarios to evaluate IK-based smartphone avatars in real time. Further research is needed to explore how smartphone AR and VR users perceive avatars differently and how these insights could shape future immersive applications. For IK-animated avatars in smartphone AR, blending pre-animated walking motions may feel unnatural. To address this, future work could integrate IK for the feet to enhance realism. Furthermore, even though there was no significant difference for visual realism or device, the ANOVA analysis for the interaction between device and visual realism almost approached significance with a $p = 0.066$. This suggests that the device type combined with the visual realism could influence the user's perceived behavioral realism of virtual characters.

Acknowledgements Our research received funding from the Thuringian Ministry for Economic Affairs, Science, and Digital Society under Grant 5575/10-5 (MetaReal) and the German Federal Ministry of Education and Research (BMBF) under Grant 16SV8922 (MULTIPARTIES) as well CO-HUMANICS (Co-Presence of Humans and Interactive Companions for Seniors) project supported by the Carl-Zeiss-Stiftung within the framework of the "Durchbrüche 2020" program.

Funding Open Access funding enabled and organized by Projekt DEAL.

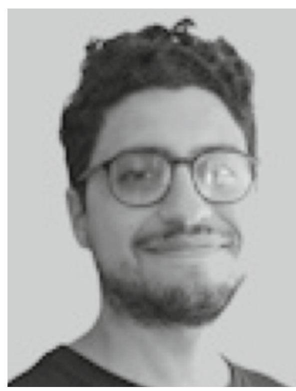
References

1. Ahuja, K., Mayer, S., Goel, M., Harrison, C.: Pose-on-the-go: approximating user pose with smartphone sensor fusion and inverse kinematics. Assoc. Comput. Mach. (2021). https://doi.org/10.1145/3411764.3445582
2. Aristidou, A., Lasenby, J.: Fabrik: a fast, iterative solver for the inverse kinematics problem. Graph. Models **73**, 243–260 (2011). https://doi.org/10.1016/j.gmod.2011.05.003
3. Aseeri, S., Interrante, V.: The influence of avatar representation on interpersonal communication in virtual social environments. IEEE Trans. Vis. Comput. Graph. **27**(5), 2608–2617 (2021). https://doi.org/10.1109/TVCG.2021.3067783

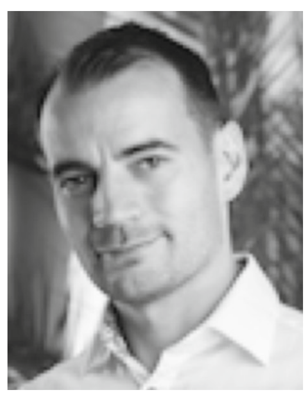
4. Bailenson, J.N., Yee, N., Merget, D., Schroeder, R.: The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. Presence: Teleoperators and Virtual Environments **15**(4), 359–372 (2006)
5. Benda, B., Ragan, E.D.: The effects of virtual avatar visibility on pointing interpretation by observers in 3d environments. In: 2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 50–59 (2021). https://doi.org/10.1109/ISMAR52148.2021.00019
6. Bevacqua, E., Richard, R., De Loor, P.: Believability and copresence in human-virtual character interaction. IEEE computer graphics and applications **37**(4), 17–29 (2017)
7. Blascovich, J., Loomis, J., Beall, A.C., Swinth, K.R., Hoyt, C.L., Bailenson, J.N.: Immersive virtual environment technology as a methodological tool for social psychology. Psychol. Inq. **13**(2), 103–124 (2002)
8. Caserman, P., Achenbach, P., Göbel, S.: Analysis of inverse kinematics solutions for full-body reconstruction in virtual reality. In: 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH), pp. 1–8 (2019). https://doi.org/10.1109/SeGAH.2019.8882429
9. Dewez, D., Hoyet, L., Lecuyer, A., Argelaguet, F.: Towards avatar-friendly 3d manipulation techniques: bridging the gap between sense of embodiment and interaction in virtual reality. Assoc. Comput. Mach. (2021). https://doi.org/10.1145/3411764.3445379
10. de Lurdes Calisto, M., Sarkar, S.: A systematic review of virtual reality in tourism and hospitality: the known and the paths to follow. Int. J. Hosp. Manag. **116**, 103623 (2024)
11. Eubanks, J.C., Moore, A.G., Fishwick, P.A., McMahan, R.P.: The effects of body tracking fidelity on embodiment of an inverse-kinematic avatar for male participants. pp. 54–63. Institute of Electrical and Electronics Engineers Inc. (2020). https://doi.org/10.1109/ISMAR50242.2020.00025
12. Everingham, M., Eslami, S.A., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes challenge: a retrospective. Int. J. Comput. Vis. **111**, 98–136 (2015)
13. Freeman, G., Zamanifard, S., Maloney, D., Adkins, A.: My body, my avatar: how people perceive their avatars in social virtual reality. Assoc. Comput. Mach. (2020). https://doi.org/10.1145/3334480.3382923
14. Gamelin, G., Chellali, A., Cheikh, S., Ricca, A., Dumas, C., Otmane, S.: Point-cloud avatars to improve spatial communication in immersive collaborative virtual environments. Pers. Ubiquitous Comput. (2021). https://doi.org/10.1007/S00779-020-01431-1
15. Gao, B., Lee, J., Tu, H., Seong, W., Kim, H.: The effects of avatar visibility on behavioral response with or without mirror-visual feedback in virtual environments. In: VR Workshops. IEEE (2020). https://doi.org/10.1109/VRW50115.2020.00241
16. Gonzalez-Franco, M., Berger, C.C.: Avatar embodiment enhances haptic confidence on the out-of-body touch illusion. IEEE Trans. Haptics **12**, 319–326 (2019). https://doi.org/10.1109/TOH.2019.2925038
17. Gonzalez-Franco, M., Ofek, E., Pan, Y., Antley, A., Steed, A., Spanlang, B., Maselli, A., Banakou, D., Pelechano, N., Orts-Escolano, S., Orvalho, V., Trutoiu, L., Wojcik, M., Sanchez-Vives, M.V., Bailenson, J., Slater, M., Lanier, J.: The rocketbox library and the utility of freely available rigged avatars. Front. Virt. Real. (2020). https://doi.org/10.3389/frvir.2020.561558
18. Groom, V., Nass, C., Chen, T., Nielsen, A., Scarborough, J.K., Robles, E.: Evaluating the effects of behavioral realism in embodied agents. Int. J. Hum. Comput. Stud. **67**, 842–849 (2009). https://doi.org/10.1016/j.ijhcs.2009.07.001
19. Herrera, F., Oh, S.Y., Bailenson, J.N.: Effect of behavioral realism on social interactions inside collaborative virtual environments. Prese. Teleop. Virt. **27**(2), 163–182 (2018). https://doi.org/10.1162/pres_a_00324
20. Jun, H., Bailenson, J.: Effects of behavioral and anthropomorphic realism on social influence with virtual humans in AR. In: 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), pp. 41–44 (2020). https://doi.org/10.1109/ISMAR-Adjunct51615.2020.00026
21. Kyrlitsias, C., Michael-Grigoriou, D.: Social interaction with agents and avatars in immersive virtual environments: A survey. Frontiers in Virtual Reality **2**, 786665 (2022)
22. Levene, H.: Robust tests of homogeneity of variance. Contributions to Probability and Statistics, Stanford Univer-sity Press, Stanford pp. 278–292 (1960)
23. Makled, E., Weidner, F., Broll, W.: Investigating User Embodiment of Inverse-Kinematic Avatars in Smartphone Augmented Reality. pp. 666–675. Institute of Electrical and Electronics Engineers Inc. (2022). https://doi.org/10.1109/ISMAR55827.2022.00084
24. Makled, E., Yassien, A., Elagroudy, P., Magdy, M., Abdennadher, S., Hamdi, N.: Pathogenius VR: VR medical training (2019). https://doi.org/10.1145/3321335.3329694
25. Makransky, G., Wismer, P., Mayer, R.E.: A gender matching effect in learning with pedagogical agents in an immersive virtual reality science simulation. J. Comput. Assist. Learn. (2019). https://doi.org/10.1111/JCAL.12335
26. Mal, D., Wolf, E., Döllinger, N., Wienrich, C., Latoschik, M.E.: The impact of avatar and environment congruence on plausibility, embodiment, presence, and the proteus effect in virtual reality. IEEE Trans. Vis. Comput. Graph. **29**(5), 2358–2368 (2023). https://doi.org/10.1109/TVCG.2023.3247089
27. Marougkas, A., Troussas, C., Krouska, A., Sgouropoulou, C.: How personalized and effective is immersive virtual reality in education? A systematic literature review for the last decade. Multimed. Tools Appl. **83**(6), 18185–18233 (2024)
28. Maselli, A., Kilteni, K., López-Moliner, J., Slater, M.: The sense of body ownership relaxes temporal constraints for multisensory integration. Sci. Rep. (2016). https://doi.org/10.1038/srep30628
29. Mauchly, J.W.: Significance test for sphericity of a normal n-variate distribution. Ann. Math. Stat. **11**(2), 204–209 (1940)
30. Milgram, P., Kishino, F.: A taxonomy of mixed reality visual displays. IEICE TRANSACTIONS on Information and Systems **77**(12), 1321–1329 (1994)
31. Mollyn, V., Arakawa, R., Goel, M., Harrison, C., Ahuja, K.: Imuposer: Full-body pose estimation using imus in phones, watches, and earbuds. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI '23. Association for Computing Machinery, New York (2023). https://doi.org/10.1145/3544548.3581392
32. Murugan, A., Vanukuru, R., Pillai, J.: Towards Avatars for Remote Communication Using Mobile Augmented Reality. pp. 135–139. Institute of Electrical and Electronics Engineers Inc. (2021). https://doi.org/10.1109/VRW52623.2021.00032
33. Peck, T.C., Gonzalez-Franco, M.: Avatar embodiment. A standardized questionnaire. Front. Virt. Real. (2021). https://doi.org/10.3389/frvir.2020.575943
34. Rogers, S.L., Broadbent, R., Brown, J., Fraser, A., Speelman, C.P.: Realistic motion avatars are the future for social interaction in virtual reality. Front. Virt. Real. (2022). https://doi.org/10.3389/frvir.2021.750729
35. Schott, E., Makled, E.B., Zoeppig, T.J., Muehlhaus, S., Weidner, F., Broll, W., Froehlich, B.: Unitexr: Joint exploration of a real-world museum and its digital twin. In: Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology, VRST '23. Association for Computing Machinery, New York (2023). https://doi.org/10.1145/3611659.3615708
36. Schwind, V., Knierim, P., Chuang, L., Henze, N.: "where's pinky?": The Effects of a Reduced Number of Fingers in Virtual Reality.

pp. 507–515. Association for Computing Machinery, Inc (2017). https://doi.org/10.1145/3116595.3116596
37. Schwind, V., Knierim, P., Tasci, C., Franczak, P., Haas, N., Henze, N.: "These are Not My Hands!": Effect of Gender on the Perception of Avatar Hands in Virtual Reality. pp. 1577–1582. Association for Computing Machinery (2017). https://doi.org/10.1145/3025453.3025602
38. Shapiro, S.S., Wilk, M.B.: An analysis of variance test for normality (complete samples). Biometrika **52**(3/4), 591–611 (1965)
39. Skarbez, R., Neyret, S., Brooks, F.P., Slater, M., Whitton, M.C.: A psychophysical experiment regarding components of the plausibility illusion. IEEE Trans. Vis. Comput. Graph. (2017). https://doi.org/10.1109/TVCG.2017.2657158
40. Steed, A., Pan, Y., Zisch, F., Steptoe, W.: The Impact of a Self-avatar on Cognitive Load in Immersive Virtual Reality. pp. 67–76. IEEE Computer Society (2016). https://doi.org/10.1109/VR.2016.7504689
41. Sterna, R., Cybulski, A., Igras-Cybulska, M., Pilarczyk, J., Segiet, N., Kuniecki, M.: How behavioral, photographic, and interactional realism influence the sense of co-presence in VR. An investigation with psychophysiological measurement. Int. J. Human–Computer Interact. (2023). https://doi.org/10.1080/10447318.2023.2285641
42. Swinth, K., Blascovich, J., et al.: Perceiving and responding to others: Human–human and human–computer social interaction in collaborative virtual environments. In: Proceedings of the 5th Annual International Workshop on PRESENCE, vol. 392. Citeseer Los Angeles (2002)
43. Thaler, M., Schlogl, S., Groth, A.: Agent Versus Avatar: Comparing Embodied Conversational Agents Concerning Characteristics of the Uncanny Valley. pp. 1–6. IEEE (2020). https://doi.org/10.1109/ICHMS49158.2020.9209539
44. Wagner, D., Billinghurst, M., Schmalstieg, D.: How Real Should Virtual Characters be? p. 57. ACM (2006). https://doi.org/10.1145/1178823.1178891
45. Wainer, J., Feil-Seifer, D.J., Shell, D.A., Mataric, M.J.: Embodiment and Human–Robot Interaction: A Task-Based Perspective. pp. 872–877. IEEE (2007). https://doi.org/10.1109/ROMAN.2007.4415207
46. Waltemate, T., Gall, D., Roth, D., Botsch, M., Latoschik, M.E.: The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. IEEE Trans. Vis. Comput. Graph. **24**, 1643–1652 (2018). https://doi.org/10.1109/TVCG.2018.2794629
47. Wang, I., Smith, J., Ruiz, J.: Exploring virtual agents for augmented reality. Assoc. Comput. Mach. (2019). https://doi.org/10.1145/3290605.3300511
48. Weidner, F., Boettcher, G., Arboleda, S.A., Diao, C., Sinani, L., Kunert, C., Gerhardt, C., Broll, W., Raake, A.: A systematic review on the visualization of avatars and agents in AR & VR displayed using head-mounted displays. IEEE Trans. Vis. Comput. Graph. **29**(5), 2596–2606 (2023)
49. Winkler, A., Won, J., Ye, Y.: Questsim: Human motion tracking from sparse sensors with simulated avatars. In: SIGGRAPH Asia 2022 Conference Papers, SA '22. Association for Computing Machinery, New York (2022). https://doi.org/10.1145/3550469.3555411
50. Yang, J.J., Chen, T., Qin, F., Lam, M.S., Landay, J.A.: Hybridtrak: Adding full-body tracking to VR using an off-the-shelf webcam. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22. Association for Computing Machinery (2022). https://doi.org/10.1145/3491102.3502045
51. Yassien, A., Elagroudy, P., Makled, E., Abdennadher, S.: A design space for social presence in VR (2020). https://doi.org/10.1145/3419249.3420112
52. Yassien, A., Makled, E., Elagroudy, P., Sadek, N., Abdennadher, S.: Give-me-a-hand: the efect of partner's gender on collaboration quality in virtual reality (2021). https://doi.org/10.1145/3411763.3451601
53. Yu, K., Gorbachev, G., Eck, U., Pankratz, F., Navab, N., Roth, D.: Avatars for teleconsultation: effects of avatar embodiment techniques on user perception in 3d asymmetric telepresence. IEEE Trans. Vis. Comput. Graph. **27**, 4129–4139 (2021). https://doi.org/10.1109/TVCG.2021.3106480
54. Zhang, L., Sturm, J., Cremers, D., Lee, D.: Real-time human motion tracking using multiple depth cameras. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2389–2395 (2012). https://doi.org/10.1109/IROS.2012.6385968
55. Zibrek, K., McDonnell, R.: Social presence and place illusion are affected by photorealism in embodied VR. In: MIG. ACM (2019). https://doi.org/10.1145/3359566.3360064

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

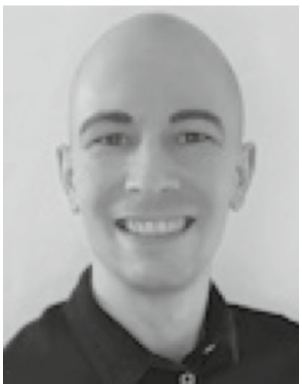


Elhassan Makled is a PhD candidate in computer science and automation and a research assistant at the Virtual Worlds and Digital Games research group at Ilmenau University of Technology. His research interests encompass avatar representation and embodiment, mediated reality, and multi-user cross-reality applications.



Christoph Gerhardt is a researcher at Ilmenau University of Technology, where he focuses on deep learning-based image processing and scene understanding. Holding a master of science degree, his research interests align with the advancements in computer vision, particularly in understanding complex scenes and interpreting visual data.

Tobias Schwandt is a research assistant at the Virtual Worlds and Digital Games group at Ilmenau University of Technology since 2014. He earned his M.Sc. in applied computer science from the University of Applied Sciences Erfurt. His research focuses on mixed reality (MR), particularly illumination in augmented reality (AR), ambient lighting reconstruction, and mediated reality. In 2021, he obtained his PhD. He has since supervised doctoral students, managed AR/VR projects, and acquired research funding.



Florian Weidner is a lecturer at the University of Glasgow, specializing in interaction in AR/VR. Previously, he was a postdoctoral researcher at Lancaster University. Before that, he was a postdoctoral researcher at Ilmenau University of Technology, where he also completed his PhD. His research examines how gaze and body movements shape immersive experiences in augmented and virtual reality.

Wolfgang Broll is a professor for Virtual Worlds and Digital Games and director of the Institute for Intelligent and Interactive Immersive Media and Technologies (I4MT) at Ilmenau University of Technology. He received his PhD from Tübingen University. He is currently concerned with collaborative XR, mediated reality, and shaping the future Metaverse.