



OPEN ACCESS

EDITED BY

Mincheol Shin,
Tilburg University, Netherlands

REVIEWED BY

Joomi Lee,
University of Arkansas, United States
Ayush Bhargava,
Meta, United States

*CORRESPONDENCE

David Mal,
✉ david.mal@uni-wuerzburg.de

RECEIVED 13 April 2024

ACCEPTED 25 June 2024

PUBLISHED 19 July 2024

CITATION

Mal D, Döllinger N, Wolf E, Wenninger S,
Botsch M, Wienrich C and Latoschik ME (2024),
Am I the odd one? Exploring (in)congruencies in
the realism of avatars and virtual others in
virtual reality.

Front. Virtual Real. 5:1417066.

doi: 10.3389/frvir.2024.1417066

COPYRIGHT

© 2024 Mal, Döllinger, Wolf, Wenninger,
Botsch, Wienrich and Latoschik. This is an open-
access article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Am I the odd one? Exploring (in)congruencies in the realism of avatars and virtual others in virtual reality

David Mal^{1*}, Nina Döllinger², Erik Wolf^{1,3}, Stephan Wenninger⁴,
Mario Botsch⁴, Carolin Wienrich² and Marc Erich Latoschik¹

¹Human-Computer Interaction Group, University of Würzburg, Würzburg, Germany, ²Psychology of Intelligent Interactive Systems Group, University of Würzburg, Würzburg, Germany, ³Human-Computer Interaction, Universität Hamburg, Hamburg, Germany, ⁴Computer Graphics Group, TU Dortmund University, Dortmund, Germany

Virtual humans play a pivotal role in social virtual environments, shaping users' VR experiences. The diversity in available options and users' individual preferences can result in a heterogeneous mix of appearances among a group of virtual humans. The resulting variety in higher-order anthropomorphic and realistic cues introduces multiple (in)congruencies, eventually impacting the plausibility of the experience. However, related work investigating the effects of being co-located with multiple virtual humans of different appearances remains limited. In this work, we consider the impact of (in)congruencies in the realism of a group of virtual humans, including co-located others (agents) and one's self-representation (self-avatar), on users' individual VR experiences. In a 2 × 3 mixed design, participants embodied either (1) a personalized realistic or (2) a customized stylized self-avatar across three consecutive VR exposures in which they were accompanied by a group of virtual others being either (1) all realistic, (2) all stylized, or (3) mixed between stylized and realistic. Our results indicate groups of virtual others of higher realism, i.e., potentially more congruent with participants' real-world experiences and expectations, were considered more human-like, increasing the feeling of co-presence and the impression of interaction possibilities. (In)congruencies concerning the homogeneity of the group did not cause considerable effects. Furthermore, our results indicate that a self-avatar's congruence with the participant's real-world experiences concerning their own physical body yielded notable benefits for virtual body ownership and self-identification for realistic personalized avatars. Notably, the incongruence between a stylized self-avatar and a group of realistic virtual others resulted in diminished ratings of self-location and self-identification. This suggests that higher-order (in)congruent visual cues that are not within the ego-central referential frame of one's (virtual) body, can have an (adverse) effect on the relationship between one's self and body. We conclude on the implications of our findings and discuss our results within current theories of VR experiences, considering (in)congruent visual cues and their impact on the perception of virtual others, self-representation, and spatial presence.

KEYWORDS

virtual human, virtual others, avatar, agent, sense of embodiment, co-presence, plausibility, presence

1 Introduction

Social virtual environments (SVEs) have gained significant attention for their remarkable ability to foster pro-social interactions and push the boundaries of traditional digital collaboration platforms (McVeigh-Schultz et al., 2019; Schulz, 2023). Within this realm, virtual humans can play a pivotal role, facilitating a multitude of mixed, augmented, and virtual reality (MR, AR, VR: XR for short) experiences. As so-called avatars, they can be directly controlled and embodied by users (Bailenson and Blascovich, 2004; Slater et al., 2009), enabling a bodily experience and sharing social signals in SVEs (Bente et al., 2008; Kolesnichenko et al., 2019; Freeman and Maloney, 2021). Yet, the versatility of virtual humans within SVEs transcends self-representation, as they may function as computer-controlled agents (Bailenson and Blascovich, 2004), or seamlessly blend into embodied ambient crowds (Latoschik et al., 2019). Numerous commercially available SVEs like VRChat, RecRoom, and Mozilla Hubs (as outlined in Liu and Steed (2021)) provide a broad spectrum of virtual humans varying in their appearances or even allow users to upload their individual avatars (Kolesnichenko et al., 2019; Hepperle et al., 2022). Therefore, the styles of virtual humans and their degree of realism and individualization can differ significantly. Realistic virtual humans may resemble life- and human-like visual features (Latoschik et al., 2017), while stylized or abstract virtual humans yield a rather simplified or iconic style (Lugrin et al., 2015; Zell et al., 2015). Individualization, on the other hand, tailors a virtual human to a specific individual, fostering the digital representation to become congruent with the user's physical appearance. Conversely, a more generalized appearance enables playful and creative avatar selection or the maintenance of a desired distance from one's physical body. Users' avatar choices can also be influenced by their access to avatar-creation technologies. While some have the ability to craft personalized, life-like avatars through 3D reconstruction from multi-view scanning (Achenbach et al., 2017) or by using smartphone cameras (Wenninger et al., 2020), others may settle for customizing virtual humans or selecting generic ones provided by the application (Hepperle et al., 2022). This diversity of options and individual preferences can result in a heterogeneous mix of various appearances among virtual humans in SVEs.

A substantial body of research has placed emphasis on understanding how the visualization of avatars and agents impacts users' virtual experiences and their evaluation of co-located virtual others (Nowak and Fox, 2018; Weidner et al., 2023). However, related work investigating the effects of being co-located with multiple virtual humans of different styles seems limited. Latoschik et al. (2019) suggested that employing mixed appearances for multiple virtual others in an ambient crowd could enhance participants' interest in the virtual environment. However, they might also increase feelings of eeriness (Latoschik et al., 2019) and introduce incongruencies within a group of virtual humans. In this context, the concepts of congruence and plausibility have emerged as fundamental when exploring VR experiences (Skarbez et al., 2021; Latoschik and Wienrich, 2022; Slater et al., 2022). This includes considering virtual humans and their

(in)congruent appearance and behavior within a particular virtual environment as an essential feature contributing to a user's plausible VR experience (Skarbez et al., 2017b; Wolf et al., 2022c; Mal et al., 2022). Related work has investigated typical cues potentially influencing an (in)congruent appearance and behavior of virtual humans, including various factors of virtual human visualization (Weidner et al., 2023) such as realism (Latoschik et al., 2017; Zibrek et al., 2019) or personalization (Waltemate et al., 2018; Fribourg et al., 2020). In turn, these cues have been shown to potentially influence various qualia reflecting the VR experience, such as spatial presence (the feeling of really being in a VE) (Slater, 2009), the sense of embodiment (the feeling of being inside, having, and controlling an avatar in a VE) (Slater et al., 2009), as well as, co-presence (the subjective experience of being in the company of virtual others) (Schroeder, 2002). Yet, to our knowledge, there has not been a further investigation into the effects of being co-located with multiple virtual humans of different styles in an avatar-mediated VR setting, considering both the realism of co-located virtual humans, the realism of the avatar, and their (in)congruencies. We conclude with the following research question:

RQ: How do the avatar's realism, co-located virtual humans' realism, and their (in)congruencies affect the perception of virtual others, the self-presentation, and the overall VR experience?

We investigate the stated research question by focusing on the intricate dynamics emerging from (in)congruent styles of a group of virtual humans, including multiple co-located others (agents), and one's digital self-representation (avatar). Therefore, we conducted a user study in which 48 participants each embodied an individualized avatar with varying degrees of realism, i.e., either being realistic or stylized, while consecutively being accompanied by three groups of virtual others also varying in their realism (i.e., all realistic, all stylized, or mixed). Realistic avatars were personalized by scanning participants with a custom-made photogrammetry rig and applying a 3D-reconstruction photogrammetry pipeline. Stylized avatars were customized by the participants using a lightweight graphical user interface. Accompanied by virtual others, each participant engaged in three VR exposures, utilizing an odd-one-out logic task paradigm. Embodiment was implemented with a state-of-the-art markerless tracking system. We evaluate the impact of the realism of a group of virtual others, the self-avatar realism, and the resulting (in)congruencies on users' VR experiences and discuss the results in the context of current theories and models, providing valuable insights for designing and developing future social virtual environments utilizing virtual humans.

2 Related work

2.1 Virtual humans

Drawing from diverse specifications in previous research (Nowak and Fox, 2018; Burden and Savin-Baden, 2019; Doerner et al., 2022; Kyrlitsias and Michael-Grigoriou, 2022), we adopt a

conceptual and operational definition of virtual humans as user (avatar) or system (agent) controlled digital representations of human beings in a virtual environment. The characteristics of virtual humans are determined by various technical components typically related to their general form (e.g., shape and resolution of mesh, dimensions, and proportions), surface (e.g., topology, shading, and texture), motion (e.g., facial and body animation), and sound (e.g., voice or heartbeat) (Lugrin et al., 2015; Burden and Savin-Baden, 2019; Zell et al., 2019), collectively contributing to the life- and human-like qualities of the virtual human's behavior and appearance.

2.1.1 Realistic and anthropomorphic cues in virtual humans

Deriving from Nowak and Fox (2018), we refer to the realism of virtual humans as the perception that they could realistically or possibly exist in a non-mediated context (life-like) and anthropomorphism as the perception or assignment of human traits or qualities to these entities (human-like). We argue that the realism of virtual humans in their literal and conceptual meaning is fundamentally rooted in anthropomorphism, as the life-likeness of a human representation depends on its contingent to incorporate human-like features. Therefore, the realism of virtual humans, as an overarching term, may be characterized by an array of realistic and anthropomorphic cues. Further specifying, Lugrin et al. (2015) discerned two categories of anthropomorphic cues in the appearance of virtual humans: anatomy, which details the structure, number, and interconnections of body parts, and composition, which describes the technical properties of specific body parts. A classification of cues can be the base for comparing and categorizing virtual humans of various artistic styles. In a recent review, Weidner et al. (2023) classified realistic visualizations of avatars and agents as detailed models of humans based on real persons and a stylized visualization to maintain human proportions with detailed body parts, yet lacking human-like textures and not necessarily adhering to human morphology. Stylization can also be classified along iconic and non-iconic scales, and the stylization of individual cues of virtual humans may also be considered independently (Zell et al., 2019), e.g., across shape (Zell et al., 2015), rendering styles (McDonnell et al., 2012; Zell et al., 2015; Volante et al., 2016; Wisessing et al., 2016; Zibrek et al., 2017), or composition (Lugrin et al., 2015). Overall, Weidner et al. (2023) indicates realism in avatars and agents to potentially benefit a multitude of qualia related to the VR experience.

2.1.2 Individualization and truthfulness of virtual humans

Virtual humans can differ not only in style but also in whether and how they are individualized, i.e., tailored to a specific individual, and the resulting truthfulness, i.e., the degree of similarity between the user's appearance and the virtual human (Gorisse et al., 2019). Thereby, we refer to personalization as the process of creating a virtual human based on user (photogrammetric) data (Waltemate et al., 2018) and to customization as the process through which a user actively alters the visual properties of a virtual human (Ducheneaut et al., 2009). The individualization of virtual humans as self-avatars can facilitate the acceptance of the virtual body as one's own and also enhance the identification with the self-

avatar, which applies for personalization (Waltemate et al., 2018; Gorisse et al., 2019; Fiedler et al., 2023; Salagean et al., 2023) as well as customization (Döllinger et al., 2023; Lee et al., 2023).

2.2 Exploring (in)congruencies in a group of virtual humans

Coherence and plausibility have gained significant attention as fundamental concepts in describing and classifying mixed, augmented, and virtual reality experiences (Skarbez et al., 2021; Latoschik and Wienrich, 2022; Slater et al., 2022). As introduced by Slater (2009), plausibility describes the illusion that what is happening is really happening. Coherence, on the other hand, refers to the extent to which a virtual scenario behaves reasonably and predictably, thus creating the illusion of plausibility (Skarbez et al., 2017a). Latoschik and Wienrich (2022) introduced an alternative theoretical Congruence and Plausibility (CaP) model. It proposes plausibility and congruence to become central conditions in describing XR experiences and effects. According to the CaP model, congruence, as the ontological specification of coherence, describes the objective match between processed and expected cues on the sensory, perceptual, and cognitive layers, creating a state of plausibility that influences various qualia and constructs of XR. With reference to the CaP model, virtual humans, whether as avatars or agents, and the congruence of their cues can also contribute to a user's plausible XR experience. In this regard, Mal et al. (2022) framed the subjective feeling of how reasonable and believable a virtual human appears to a user as *virtual human plausibility (VHP)*. VHP would, therefore, arise from the congruence of habitual sensory, proximal perceptual, or higher-order cognitive cues of the virtual human in the VE and eventually impact various qualia shaping the XR experience. In this work, we explore (in)congruent appearances in a group of virtual humans, including co-located others (agents) and one's self-representation (self-avatar). Therefore, following the CaP model, we will subsequently identify manipulated cues, delineate their (in) congruencies, and pinpoint relevant qualia potentially influenced by the specific manipulation as summarized in Table 1.

2.2.1 Manipulation space

We classify our manipulation of appearance on a common scale between realism and stylization (Zell et al., 2019). Therefore, we chose two types of virtual humans distinct in cues related to their form and texture:

- (1) Realistic virtual humans created by a 3D-reconstruction photogrammetry process striving for life and human-like appearance.
- (2) Stylized virtual humans of lower realism and anthropomorphism, based on a cartoon-style 3D model with human anatomy but simplified composition.

As for the group of others, we manipulated the group's configuration to be either realistic or stylized, i.e., a homogeneous group of virtual humans, or mixed, i.e., a group evenly distributed between stylized and realistic virtual humans. Furthermore, self-avatars were either realistic, based on scan-based personalization, or

TABLE 1 Overview specifying our manipulation and related cues for virtual others and the self-avatar in a group of virtual humans, further naming expected (in)congruencies, and defining this works' qualia space.

	Manipulation		(In)Congruencies	Qualia
	Specification	Related cues		
Virtual Others	Realistic, Mixed, Stylized	Realism Anthropomorphism	Lifelikeness Homogeneity Self-Avatar	Virtual Human Plausibility Co-Presence Affective Appraisal Spatial Presence
Self-Avatar	Realistic, Stylized	Realism Anthropomorphism Truthfulness	Physical Body Virtual Others	Sense of Embodiment Self-Identification Spatial Presence

TABLE 2 Descriptive statistics of our measures assessed after each VR exposure compared for each self-avatar and the three configurations of virtual others.

Measures	Range	Stylized others		Mixed others		Realistic others	
		Stylized self	Realistic self	Stylized self	Realistic self	Stylized self	Realistic self
		<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Virtual Human Plausibility							
Appearance and Behavior	[1–7]	4.73 (1.29)	4.56 (1.30)	4.55 (0.95)	4.49 (1.17)	4.99 (0.99)	4.89 (1.15)
Match with VE	[1–7]	4.79 (1.43)	4.25 (1.60)	4.24 (1.30)	4.46 (1.52)	5.09 (1.37)	5.27 (1.15)
Co-Presence							
Co-Presence	[1–7]	4.29 (1.49)	3.58 (1.44)	4.97 (1.29)	4.50 (1.21)	5.38 (1.15)	5.03 (1.24)
Interaction Possibilities	[1–7]	2.28 (1.13)	1.97 (0.96)	2.73 (1.17)	2.70 (1.12)	3.25 (1.51)	3.06 (1.68)
Affective Appraisal							
Humanness	[1–7]	2.49 (0.96)	2.23 (0.95)	3.37 (0.95)	3.57 (1.26)	4.28 (1.37)	4.23 (1.19)
Eeriness	[1–7]	2.99 (1.05)	3.34 (1.37)	3.30 (1.01)	3.67 (1.05)	3.48 (1.00)	3.33 (0.81)
Sense of Embodiment							
Virtual Body Ownership	[1–10]	5.08 (2.12)	6.21 (1.74)	4.75 (2.29)	6.38 (1.84)	4.50 (2.30)	6.54 (1.79)
Agency	[1–10]	6.29 (1.88)	6.79 (1.82)	6.08 (2.08)	6.79 (1.47)	5.83 (1.79)	7.00 (1.56)
Change	[1–10]	4.04 (3.38)	4.50 (2.72)	3.92 (3.31)	4.67 (2.73)	3.71 (3.04)	4.71 (2.69)
Self-Location	[1–10]	5.83 (2.08)	5.42 (2.26)	5.21 (2.36)	5.62 (2.14)	4.83 (2.60)	5.67 (2.12)
Self-Identification							
Self-Identification	[1–10]	5.02 (2.03)	6.75 (1.68)	4.48 (1.90)	7.08 (1.04)	3.96 (1.77)	6.96 (1.46)
Spatial Presence							
Spatial Presence	[1–10]	6.46 (2.00)	6.33 (1.27)	6.42 (1.44)	6.88 (1.48)	6.33 (2.12)	7.00 (1.47)

stylized, based on customization. We assume realistic avatars to be of a higher level of truthfulness as they objectively resemble more visual features similar to the user's real appearance than the stylized ones. Within the technical system's boundaries (see Section 3.4.3), the scan-based avatars are created using users' photogrammetric data, depicting the user in form and texture. On the other hand, customized stylized avatars are constrained by available customization options and a simplified composition.

2.2.2 Condition space

Drawing from the CaP model, we hypothesize the named manipulation of realism across a group of virtual humans to result in multiple (in)congruencies in higher-order visual cues on a cognitive layer.

- (1) (In)congruencies within the group of virtual others depending on its style configuration. Overall, a homogeneous group configuration appears more congruent than a mixed one, an (in)congruence that can be accessed by directly comparing others within the VE. Furthermore, we assume the realism of each group member to shape the entire group's congruence with the participant's real-world experiences and expectations. With the highest realism/congruence for realistic others, less for mixed, and lowest for stylized others.
- (2) The self-avatar's (in)congruence with the participant's real-world experiences concerning their own physical body. We expect a personalized realistic avatar to be of higher realism and truthfulness and, therefore, more congruent with

participants' experiences and expectations towards their physical body than a customized stylized avatar.

- (3) The group of virtual others' (in)congruence with the self-avatar, and *vice versa*. We suggest the self-avatar to be congruent with a homogeneous group of the same style, less congruent with mixed groups, and incongruent with the homogeneous group of the opposite style.

2.2.3 Qualia space

While the CaP model defines the overall frame for manipulating cues and their congruencies, it does not provide a ranking of (in)congruencies or specify changes in the qualia space. Therefore, in the following, we identify relevant qualia and constructs describing the VR experience and, thereby, review related research indicating how the named (in)congruencies may influence the perception of co-located virtual others, the perception of self-representation, and the VR experience.

2.3 (In)congruencies and the perception of virtual others

Prior research has examined how the congruence in various sensory impressions of virtual others affects the XR experience. The work of Skarbez et al. (2017b) investigated a virtual agent's behavior coherence and its relative importance as a contributing factor for the overall plausibility of a VR experience. Further research focused on the congruence of spatial and behavioral cues of agents (Kim et al., 2017), gaze behavior, and auditory features of virtual groups (Bergström et al., 2017), facial animation methods (Kullmann et al., 2023), different virtual body animation features for avatars and agents (Debarba et al., 2022), and renderings of single virtual humans and their (in)congruence with the device-related presentation of the respective environment (Wolf et al., 2022c). However, we are unaware of related studies investigating the congruence of styles within a group of virtual humans. Following Mal et al. (2022)'s conceptualization of VHP, (in)congruencies in the realism of virtual others and their (in)congruence with the self-avatar might impact the perceived plausibility of virtual others'. We hypothesize more congruent conditions to result in a higher attribution of plausibility towards others and deduce the following hypotheses.

- H1: The manipulation of virtual others' realism and the self-avatar's realism will lead to significantly higher scores in VHP for configurations of higher congruence.

2.3.1 Co-presence and the impression of interaction possibilities

Co-presence describes the subjective experience of being in the company of others in a virtual environment, or in short, a sense of "being there together" (Schroeder, 2002). We refer to it as a quale, denoting the sensation of being together in a (virtual) place (Skarbez et al., 2017a). As noted by Kyriltsias and Michael-Grigoriou (2022), co-presence (referred to as social presence by the authors) can enhance the realism and effectiveness of interactions between users and virtual humans. One decisive

feature influencing co-presence in virtual environments is the visual representation of others. While the congruence between realism in appearance and behavior has been named of great importance for co-presence (Garau et al., 2003; Bailenson et al., 2005), previous work provided mixed results about the significance of a realistic or anthropomorphic appearance of avatars and agents (Oh et al., 2018). For example, in the work of Latoschik et al. (2017), the realism of avatars (abstract/realistic) in dyadic VR scenarios did not affect co-presence, though the authors reported slightly higher eeriness ratings for the realistic avatars potentially indicating an Uncanny Valley effect. In contrast, Zibrek et al. (2017) found an impact of the rendering styles of animated agents on co-presence in two VR experiments. Interestingly, realistic virtual humans were preferred over stylized representations, while the rendering styles were rated comparable in unappealing and eeriness. Also, Volante et al. (2016) indicated interactions with virtual patients of realistic rendering to increase emotional bonding and social presence compared to stylized ones. For multiple virtual others, Latoschik et al. (2019) indicated the impression of interaction possibilities (IIP) and co-presence to be consistently higher for a mixed ambient crowd of abstract and realistic virtual participants' along with increased feelings of eeriness. The authors reason that potential incongruencies with participants' expectations may have caused them to focus more intensely on the surrounding agents when they were of mixed appearance. Lastly, the realism in shape and texture of the self-representation and its congruence with the appearance of an agent was not found to impact co-presence in the work of Latoschik et al. (2017). We did not find further indication for the (in)congruence in realism between the self-avatar and others to impact co-presence and IIP. In conclusion, we acknowledge that related work on co-presence, specifically investigating the effects of being co-located with multiple virtual humans of different styles in immersive VR, is limited and rather inconclusive. Therefore, we concur with the systematic review by Oh et al. (2018), which affirms mixed results regarding the impact of others' realism on co-presence. We propose an exploratory evaluation of how our manipulation affects co-presence and IIP.

2.3.2 Affective appraisal and the uncanny valley effect

As indicated, incongruencies in realistic and anthropomorphic cues of virtual humans might lead to increased feelings of eeriness or unappealing towards virtual humans. In this regard, the Uncanny Valley effect delineates a phenomenon describing a transition from an initial affinity to a feeling of eeriness as the appearance of a robot or any anthropomorphic character approaches a convincingly human-like representation but falls short of achieving it (Mori et al., 2012). There have been indications of an uncanny valley effect in VR for anthropomorphic virtual characters of different visual realism or humanness (Latoschik et al., 2017; Hepperle et al., 2022). To control for the potential impact of an Uncanny Valley effect, or a general sense of discomfort towards virtual others, we propose an exploratory evaluation of how their eeriness is perceived. Additionally, we aim to assess our manipulation in terms of perceived humanness.

2.4 (In)congruencies and the perception of self-representation

2.4.1 Sense of embodiment and self-identification

An essential quale describing the experience of having an avatar is the sense of embodiment (SoE). It emerges when a virtual body's properties are processed as if they were the properties of one's own physical body (Kilteni et al., 2012). SoE has been named an essential concept for other effects like the *Proteus* effect to emerge (Ratan et al., 2019; Mal et al., 2023). It is considered to consist of three sub-dimensions describing the senses of having (body ownership), controlling (agency), and being inside (self-location) a virtual body (Longo et al., 2008; Slater et al., 2010; Kilteni et al., 2012). Roth and Latoschik (2020) further named a perceived change in the body schema as an essential component of the SoE, a factor that is rather significant for studies using altered body appearances, according to the authors. The named sub-components can be assessed with well-established questionnaires (Roth and Latoschik, 2020; Peck and Gonzalez-Franco, 2021). In general, the SoE is understood to stem from integrating both bottom-up and top-down influences (Kilteni et al., 2012; Maselli and Slater, 2013). Bottom-up information arises from the congruence of visual, tactile, and proprioceptive cues, predominantly determined by the capabilities of the technical system (Kilteni et al., 2012; Slater et al., 2022). On the other hand, top-down information is derived from the cognitive processing of the avatar's visual cues and their congruencies, encompassing factors like self-similarity (Waltemate et al., 2018; Fiedler et al., 2023; Salagean et al., 2023) and realism (Latoschik et al., 2017; Salagean et al., 2023).

Regarding the visualization of avatars, a recent review by Weidner et al. (2023) indicated that virtual body ownership (VBO) indeed benefits from both personalized and realistic appearances. This suggests that congruence between users' real-world experiences, particularly with their own physical bodies, and the self-avatar can enhance the sense of owning and accepting a virtual body. For our manipulation, we expect (personalized) realistic self-avatars to be more congruent with the participant's physical appearance, potentially leading to higher VBO values. Conversely, a current meta-analysis by Mottelson et al. (2023) pointed out that the appearance of the self-avatar seems to be of less importance for the perceived agency over the virtual body. We concur with this conclusion, provided that the underlying technical system ensures comparable contingencies for visuomotor congruence across the self-avatar types, a condition we also assume to be met by our embodiment system. For self-location, we follow Kilteni et al. (2012), naming it to be primarily influenced by the visuospatial perspective, another factor we kept constant for all conditions in our system configuration.

Another important component concerning the perception of self-avatars is the process of identifying with the digital representation, namely, self-identification (González-Franco et al., 2020). Interestingly, the body of research focused on self-identification remains limited (Fiedler et al., 2023), despite its recognized significance as a pivotal factor driving relevant XR effects, like the *Proteus* effect (Ratan et al., 2019). Following Wolf et al. (2022a), self-identification encompasses two key components: self-similarity as the perceived visual congruence between the individual and the virtual human, and self-attribution as the

attribution of personal characteristics to the virtual human for both external body features or internal character traits. Salagean et al. (2023) recently investigated self-identification towards virtual humans with a full-body embodiment system, reporting higher realism and personalization to be beneficial for self-identification. We attribute these findings to higher-order visual cues determining the congruence between the participant's physical body and the self-avatar.

We conclude with the following hypotheses for the sense of embodiment and self-identification.

- H2.1: The manipulation of the self-avatar's congruence with the participant will lead to significantly higher scores in VBO and self-identification for realistic self-avatars.
- H2.2: The manipulation of the self-avatar's congruence with the participant will not significantly affect agency, change, and self-location.

2.4.2 The perception of self-representation and virtual others

Interestingly, Latoschik et al. (2017) found a marginal effect of an agent's realism (realistic vs abstract) on embodied users' change in self-perceived body schema (change). This suggests that the appearance of virtual others can extend beyond affecting general user experience or the evaluation of the virtual others. It may also influence self-related concepts within the ego-central referential frame of one's (virtual) body, such as the SoE and self-identification toward an avatar. It is worth noting that insight in exploring these concepts dependent on the visualization of virtual others remains limited (Mal, 2020), especially in multi-agent embodied virtual environments. We, therefore, propose an exploratory evaluation of the impact of the realism of virtual others and the congruence between the self-avatar and virtual others on the perception of the self-avatar.

2.5 Spatial presence

Spatial presence is an essential concept in evaluating a user's individual VR experience and can be considered an elementary foundation for other VR potentials to become effective (Wienrich et al., 2021). The process model of the formation of spatial presence experiences names self-location as the core dimension of spatial presence, i.e., "the sensation of being physically situated within the spatial environment portrayed by the medium" (Wirth et al., 2007, p. 497). However, to avoid ambiguities, we adhere to the definition of Kilteni et al. (2012) and name the term self-location to refer to one's spatial experience of being inside a (virtual) body rather than being inside a world, whereas we apply the term spatial presence to the (psychological) sense of "being there" in a virtual environment (Slater and Wilbur, 1997). Spatial presence has been named to be predominantly bottom-up driven by the objective concept of immersion, which describes the capabilities of the system providing the boundaries within which spatial presence occurs (Slater et al., 2022). Latoschik and Wienrich (2022), on the other hand, assumes spatial presence to be affected by the congruence and plausibility of spatial cues on the sensation, perception, and cognition levels, not limiting influencing factors to the system's

capabilities. Concerning the avatar's congruence with the physical appearance of the participant, previous work has shown that spatial presence can benefit from personalization (Waltemate et al., 2018). However, the recent work of Salagean et al. (2023) did not reveal a significant effect of avatar personalization (high/low), realism (high/low), and their interaction on presence. As these results contradicted the authors' hypotheses, as well as the results of previous work (Waltemate et al., 2018), the authors attribute their findings to a constant level of agency and appearance of the virtual environment across all conditions. Also, Latoschik et al. (2017) found no impact of virtual human realism on spatial presence, applying to avatars and agents. Further related work indicates no effect of agent visualization on spatial presence (Rzayev et al., 2019; Butz et al., 2022).

In our experiment, we maintained a constant degree of immersion (bottom-up) across conditions. Furthermore, as argued by Salagean et al. (2023), we expect a consistent level of agency and a stable appearance of the virtual environment. Additionally, we anticipate no significant manipulation of higher-order spatial cues through the manipulation of virtual humans' realism and their (in)congruencies. Finally, we are not aware of any work indicating that the realism of virtual others significantly impacts spatial presence. Consequently, we propose the following null hypothesis.

H3: The manipulation of the self-avatar's realism, the virtual others' realism, and their congruence will not significantly affect spatial presence.

We conducted a controlled user study to investigate the stated hypotheses and proposed exploratory evaluations. Therefore, we systematically manipulated the realism of a group of virtual others and the self-avatars and evaluated participants' ratings on their VR experience based on their perception of the virtual others, self-representation, and spatial presence.

3 Materials and methods

3.1 Participants

A total of 53 undergraduate students from the University of Würzburg were recruited through a participant management system and received course credit for participation according to the study duration. We excluded data from one participant reporting uncorrected visual impairment and another reporting color blindness. Three data sets were further excluded due to technical issues. For the resulting 48 valid data sets, ages ranged from 18 to 27 years ($M = 21.73$, $SD = 2.17$), with 24 participants identifying as female and 24 as male. Participants were evenly distributed between conditions, leading to gender-balanced groups of equal sample sizes. All participants were German native speakers. One participant was new to VR, 42 participants reported having less than 10 h of VR experience, and five participants reported having over 10 h of VR experience. The study was conducted according to the Declaration of Helsinki and approved by the ethics committee of the Human-Computer-Media Institute of the University of Würzburg.

3.2 Design

We employed a 2×3 mixed design with the independent variables *self-avatar* (realistic and stylized) as the between-subject factor and *virtual others* (realistic, stylized, and mixed) as the within-subject factor. Therefore, participants either embodied a personalized realistic avatar or a customized stylized avatar in three consecutive VR exposures. Each time, they solved a set of logical tasks in VR accompanied by a group of virtual others varying in their realism, i.e., all realistic, all stylized, or evenly distributed between stylized and realistic. The within-subject manipulation was presented in counterbalanced order. Dependent variables assessed the perception of the virtual others after each VR exposure regarding plausibility, affective appraisal, and perceived co-presence and interaction possibilities. We evaluated the individuals' VR experience by assessing spatial presence. Further, we considered the perception of self-presentation in terms of the sense of embodiment and self-identification with the self-avatar after each exposure and once overall after the last exposure.

3.3 Measures

3.3.1 Plausibility of virtual others

We assessed the virtual others' plausibility with the *Virtual Human Plausibility Questionnaire* (VHPQ) (Mal et al., 2022). It consists of two dimensions: (1) The virtual human's appearance and behavior plausibility (ABP) and (2) the virtual humans' match to the VE (MVE). The 11 items were rated on a seven-point Likert scale (7 = *highest plausibility*) and assessed out of VR after each VR exposure. The questions were adapted to address multiple virtual others instead of a singular virtual human.

3.3.2 Co-presence (CoP) and impression of interaction possibilities (IIP)

We assessed co-presence (CoP) and the impression of interaction possibilities (IIP) with the equivalent sub-scales of the questionnaire proposed by Poeschl and Doering (2015). The seven items were rated on a seven-point Likert scale (7 = *highest CoP or IIP*) and assessed out of VR after each VR exposure.

3.3.3 Affective appraisal

We assessed the virtual other's humanness (UVI-H) and eeriness (UVI-E) with the equivalent dimensions of the revised *Uncanny Valley Index* (UVI) (Ho and MacDorman, 2017). The 14 items were rated using semantic differentials ranging from -3 to 3. The results were mapped to a scale of one-7 (7 = *highest UVI-H and UVI-E*). The UVI's task introduction was adapted to refer to multiple virtual others instead of a singular virtual character. The items were assessed out of VR after each VR exposure.

3.3.4 Spatial presence (SP)

We assessed spatial presence (SP) with the SP sub-dimension of the *Igroup Presence Questionnaire* (IPQ) (Schubert et al., 2001). The five items are rated on a seven-point Likert scale (6 = *highest SP*) and were assessed out of VR after the last VR exposure. We further assessed the *One Item Presence Scale* (OIPS) (Bouchard et al., 2004)

in VR after each VR exposure. The OIPS comprises one item rated on a 10-point scale ($10 = \text{highest SP}$).

3.3.5 Sense of embodiment (SoE)

We assessed the SoE and its dimensions virtual body ownership (VBO), agency (AG), and the change (CH) in the body schema using the *Virtual Embodiment Questionnaire (VEQ)* (Roth and Latoschik, 2020). Additionally, we measured self-location (SL) using the items introduced by Fiedler et al. (2023). Each factor is evaluated with four items rated on a seven-point Likert scale ($7 = \text{highest SoE}$) and was assessed out of VR after the last VR exposure. We further assessed one significant item of each dimension in VR after each VR exposure, rated on a 10-point scale ($10 = \text{highest SoE}$).

3.3.6 Self-identification (SI)

We assessed self-identification (SI) with eight items concerning self-attribution and self-similarity as introduced by Fiedler et al. (2023). The items were rated on a seven-point Likert scale ($7 = \text{highest SI}$), and were assessed out of VR after the last VR exposure. We further assessed one significant item for self-attribution and one for self-similarity in VR after each VR exposure, rated on a 10-point scale ($10 = \text{highest SI}$).

3.3.7 Control measures

We controlled for interindividual differences between participants by considering their VR experience and perceived signs of simulator sickness. Therefore, we assessed VR experience before the first VR exposure as a categorical variable measuring viewing time in hours, rated in four categories ($0, 1 \text{ to } 10, 10 \text{ to } 50, \text{ and over } 50 \text{ h}$). Furthermore, we assessed simulator sickness before the first and after the last VR exposure using the *Simulator Sickness Questionnaire* (Kennedy et al., 1993). It consists of 32 items capturing symptoms associated with simulator sickness. The total score ranges from 0 to 235.62 ($235.62 = \text{strongest}$). We consider the change between the pre and post-VR assessments. An increase in score indicates the occurrence of simulator sickness due to VR usage.

3.4 Apparatus

3.4.1 Hard- and software

We implemented the VR application using Unity in version 2020.3.25f1 (Unity Technologies, 2020). A Valve Index HMD (Valve Corporation, 2022) and two Valve Index controllers were integrated using SteamVR (Valve Corporation, 2021) and the corresponding Unity plug-in in version 2.6.1¹ (see Figure 2C). The HMD has a display resolution of 1440×1600 pixels per eye, and a total field of view of $114.1 \times 109.4^\circ$ (Wolf et al., 2022a). It ran at a refresh rate of 120 Hz. The tracking area ($3 \text{ m} \times 4 \text{ m}$) was set up with four SteamVR Base Stations 2.0. We carefully routed the HMD's cable leading to a high-end workstation (NVIDIA GeForce RTX 2080 Ti, 32 GB RAM, Intel Core i7-9700K CPU, Windows 10). The participants' fingers were tracked via the controllers' proximity sensors. We did not provide eye and facial expression tracking.

For body tracking, we employed the markerless tracking system from Capture (Stoll et al., 2011) (see Figure 2D). Movements were tracked at a rate of 100 Hz with eight FLIR Blackfly S BFS-PGE-16S2C RGB cameras connected to a high-end workstation (NVIDIA GeForce RTX 3080 Ti, 32 GB RAM, AMD Ryzen 959,00x, Ubuntu 20.04.6 LTS) using two four-port 1 GBit/s ethernet frame-grabber. The workstation ran Capture Live in version 254b (The Capture, 2021) streaming the body pose to the VR system integrating Capture's Unity plug-in². Questionnaires were assessed on a desktop PC using LimeSurvey 4.5 (LimeSurvey GmbH, 2020).

3.4.2 Virtual environment

The virtual environment featured a wooden floor enclosed by white walls accented with black elements, creating a rather neutral and minimalist aesthetic (see Figure 1). Semi-transparent footprints indicated the participant's starting position and a door to the wall behind the participant aided orientation within the space, representing a potential exit. The spacious virtual environment measured $12 \times 12 \text{ m}$, providing ample room for body movement tasks, movement towards the group of virtual others, and participation in the odd-one-out paradigm within a group of seven virtual humans. A black framed mirror ($1 \times 2 \text{ m}$) was initially hidden and later on appeared 2 m in front of the participant (see Figures 2E,F).

3.4.3 Creation of realistic virtual humans

We employed a custom-made photogrammetry rig to scan participants in a laboratory at the University of Würzburg. The rig featured 15 Canon EOS 2000D cameras equipped with high-quality 35 mm wide angle lenses³ arranged in a 5×3 grid (see Figure 2A). The small number of cameras and the streamlined design ensured unhindered accessibility to the scanning process, even within the confines of a small laboratory. Six LED tubes were used to light the subjects uniformly. To capture the participants from every angle, they were scanned from the front, back, left, and right. An additional scan captured the empty background, which was used for background subtraction in the photogrammetry step. The images were automatically processed using the commercial photogrammetry software Agisoft Metashape (Agisoft, 2021), resulting in four dense point clouds, each defining a partial scan of the subject.

For generating realistic virtual humans from this input data, we use a template-fitting approach based on the work of Achenbach et al. (2017). They used a photogrammetry rig to generate a single dense point cloud of the subject, to which an animatable statistical human body model was fitted. Non-rigid ICP (Bouaziz et al., 2014) and fine-scale surface deformation were employed to optimize the statistical template model's alignment, pose, and shape.

With our approach, participants turned 90° for each of the scans, and the four point clouds depict the subjects in slightly varying poses. We adopted the avatar generation method of Achenbach et al. (2017) accordingly. On each point cloud, we selected k landmarks, which served as initial correspondences between the template model

1 <https://assetstore.unity.com/packages/tools/integration/steamvr-plugin-32647>

2 <https://capture.com/resources/>

3 <https://www.canon.de/lenses/ef-35mm-f-2-is-usm-lens/>



FIGURE 1
A participant embodying a realistic self-avatar (front) while solving a logic task in a group of virtual others (back). The group features a configuration of mixed styles consisting of three realistic virtual humans and three stylized virtual humans.

and the partial scans ($k \in \{23, 9, 6, 6\}$ for the front, back, left, and right scans, respectively). Pose parameters of the template model were then individually optimized for each of the partial scans, while the shape parameters were jointly optimized such that the resulting shape matched all of the point clouds. Similarly, the fine-scale deformation was reformulated to minimize the shape difference between the template model and all of the partial scans simultaneously. Texture information for the resulting geometry was then generated by following the method of Wenninger et al. (2020), which generated partial textures from the camera calibration data resulting from the photogrammetry step. Partial textures were then stitched together via a graph-cut based optimization and Poisson Image Editing (Boykov et al., 2001; Pérez et al., 2003).

To improve the scan result, participants were instructed to wear close-fitting casual attire that was neither completely white nor completely black. They were also advised against wearing shoes with heels, skirts, or dresses and were recommended to choose trousers and tops made from non-reflective materials.

3.4.4 Creation of stylized virtual humans

The stylized virtual humans are based on a female⁴ and a male⁵ character 3D model in a cartoon style, both wearing long-sleeved clothing. The models were rigged using Mixamo (Adobe Systems Inc., 2023). Additional hairstyles were created using Blender in version 2.8.0 (Blender Foundation, 2019). To create individualized stylized self-avatars, participants were provided an array of customization options using a graphical user interface implemented using Unity (Unity Technologies, 2020) as depicted in Figure 2B. The options included three short and two long hairstyles, ten skin colors based on the Monk Skin Tone Scale (Monk, 2019), thirteen eye colors with variations of blue, green, brown, grey, and black, as well as thirteen hair colors, ranging from light gray, blond, red, and brown to black. A selection of 33 colors for

clothing and shoes was provided. Participants were instructed to choose a configuration that looked similar to them.

3.4.5 Implementation of the self-avatar

Participants could observe their self-avatar's virtual body from the first-person and third-person perspectives in a virtual mirror (see Figures 2E,F). We implemented embodiment by retargeting Captury's tracked body pose to the avatar in real-time. We addressed inaccuracies in the end-effectors' positions caused by variations in skeletal structure and segment lengths between the target avatar and the source pose following the approach of Wolf et al. (2022b). Thereby, we aligned the avatar's end-effectors (i.e., head, hands, and feet) with the tracked end-effectors of the user implementing an IK-supported pose optimization step using FinalIK in version 2.1 (Rootmotion, 2019). The system's motion-to-photon latency for full embodiment averaged 35 ms for hand and 116 ms for other body movements. It was assessed by frame counting (Stauffert et al., 2020), tracking the time difference between real-world movements and the rendered corresponding avatar movements. Stylized self-avatars were automatically scaled to the participant's eye height, while the personalized scan process implicitly predetermined the height of realistic self-avatars.

3.4.6 Implementation of the group of virtual others

The group of virtual others consisted of six virtual humans with three possible configurations of appearance. The virtual humans were either (1) all realistic, (2) all stylized, or (3) evenly distributed between stylized and realistic. The virtual humans' genders were distributed evenly between males and females. All virtual humans were positioned in a circular pattern with a radius of 2.5 m, facing the group's center. There was one additional empty spot for the participant to join the group (see Figure 1).

3.4.6.1 Appearance of group members

We generated six realistic virtual humans by scanning three male and three female volunteers dressed in casual attire in a laboratory at the University of Würzburg. The volunteers' ages ranged from 21 to 25 ($M = 23.33, SD = 1.86$), they gave

⁴ <https://www.cgtrader.com/3d-models/character/woman/cartoongirl009-girl>

⁵ <https://www.cgtrader.com/3d-models/character/man/cartoonman035-man>

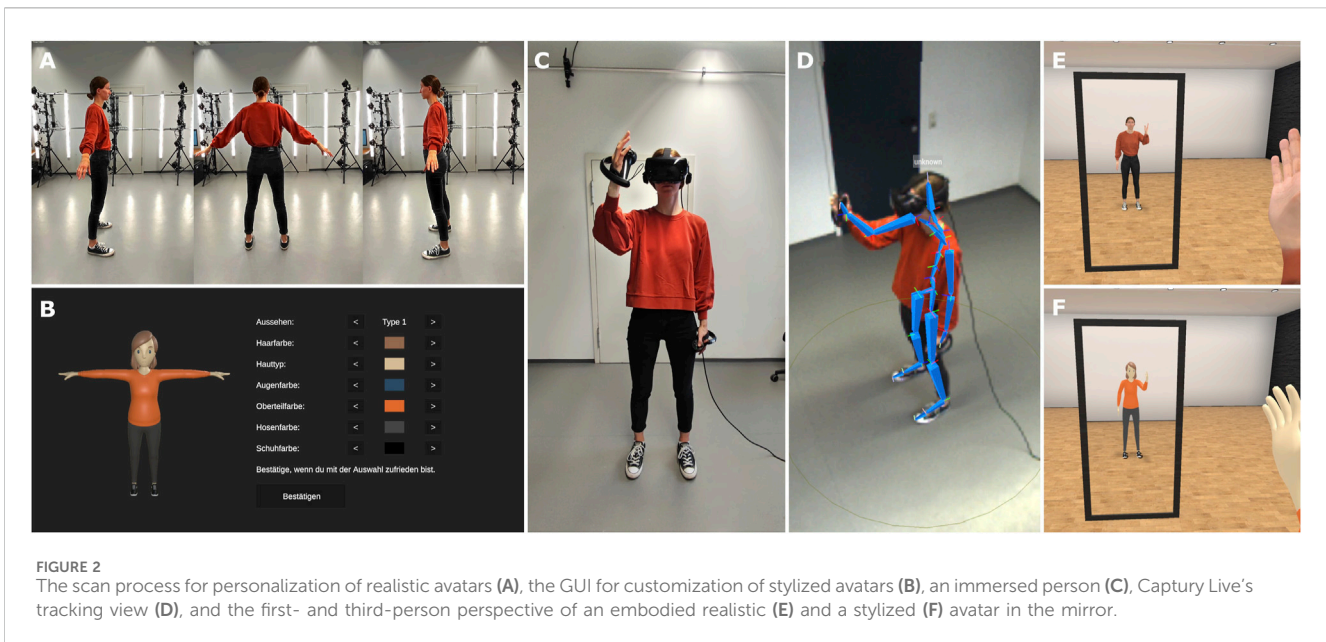


FIGURE 2 The scan process for personalization of realistic avatars (A), the GUI for customization of stylized avatars (B), an immersed person (C), Capture Live's tracking view (D), and the first- and third-person perspective of an embodied realistic (E) and a stylized (F) avatar in the mirror.

explicit consent for using their virtual reconstructions in a scientific context, and they were not compensated for participation. Three of the resulting realistic virtual humans are depicted in Figure 1. The stylized virtual others were created using random configurations of the customization options described in Section 3.4.4. Therefore, we excluded the base avatar type that participants initially chose for their customized stylized self-avatar to avoid unintentionally creating a digital twin of the participant. Stylized virtual others were scaled to a body height of 165.8 cm for females and 178.9 cm for males, relating to the body height average of the German population in 2021⁶.

3.4.6.2 Animation and behavior of group members

To avoid potential bias caused by different emotional dispositions of the virtual others (Volonte et al., 2020), we aimed to keep animation and behavior constant between the groups' configurations. Therefore, the following description of animation and behavior equally applies to both types of virtual humans. The virtual human's body animation was drawn from eight distinct idle animations exported from Mixamo (Adobe Systems Inc., 2023). To simulate the input behavior of virtual others, we additionally captured nine pointing animations using an Xsens MVN Link⁷ motion capture suit with Xsens MVN Record in version 2022.0.0 (Movella Inc., 2022). Two of the pointing gestures were performed with the left hand and seven with the right hand. All animations maintained fixed world positions and rotations, ensuring the virtual humans remained stationary, facing the group's center. No facial expressions were displayed.

The virtual others' behavior followed a rule-based, event-driven approach. Accordingly, each virtual human's behavior adhered to a currently assigned state, determining its body animation and target of attention, i.e., where they looked at in the virtual environment. All transitions between states aligned with the ongoing instructions provided to the participants, thereby yielding credible behavior that mirrored the study's procedural structure. In the default state, the virtual humans would either remain inactive or look around, performing idle animations, thereby gazing at the group center or occasionally shifting their target of attention to the eyes of other virtual humans or the participant. In an introductory state, the virtual humans' target of attention switched to the participant's eyes once they came closer than 3.5 m to the group center, which was forced when the participant was told to join the group. All virtual others then briefly looked at the participants to simulate that they recognized them. Once the logic tasks began, the virtual humans entered an operational state, directing their attention to the task objects or the input tablet respectively. Upon the participant being prompted to select an object, the virtual humans shifted state, eliciting a pointing animation towards a random solution input on the tablet.

3.4.6.3 Framing the group of virtual others

There was no explicit framing regarding whether the virtual others were human or computer-controlled. In all questionnaires, they were referred to as "virtual characters", deliberately avoiding implications of humanness, which the term "virtual humans" might have. In VR, virtual others were simply referred to as "others", and participants were instructed not to communicate with them, as each individual was required to solve the logic tasks independently. This also suggested participants not expect virtual others to initiate conversation. It was implied that the others were engaged in solving the logic tasks as well, as the audio instructions, presented while the others were also displayed, addressed

⁶ <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Gesundheitszustand-Relevantes-Verhalten/Tabellen/liste-koerpermasse.html>

⁷ <https://www.movella.com/products/motion-capture/xsens-mvn-link>

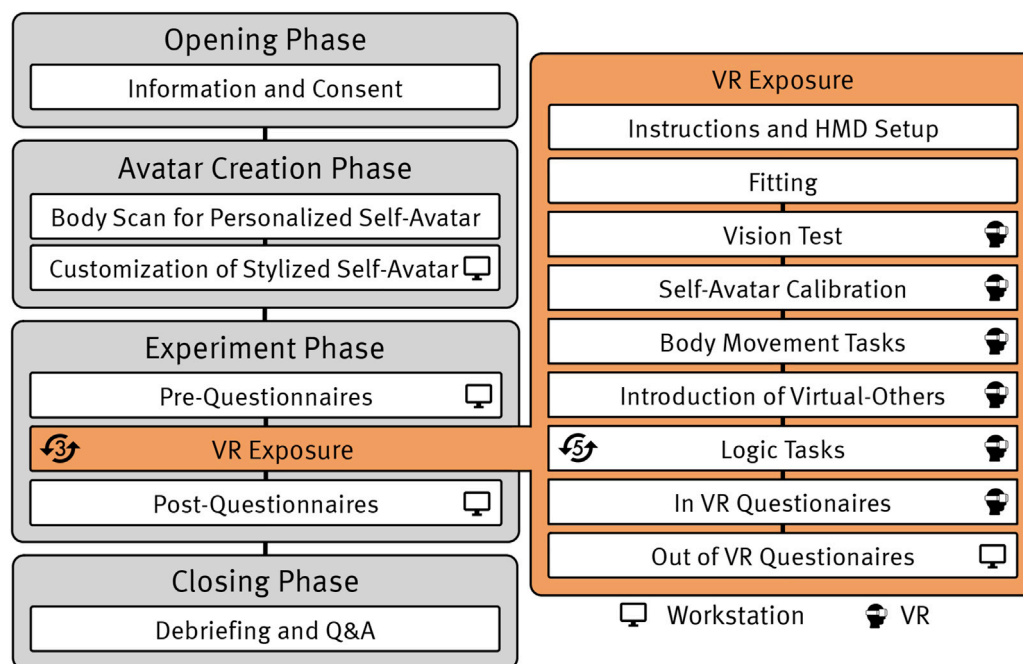


FIGURE 3

Overview of the experimental procedure (left) and the repeated VR exposures (right). All participants underwent the same procedure regardless of their randomly assigned test condition. The icons on the right side of the steps indicate whether the participants used a workstation or were immersed in VR. The icons on the left side of the steps indicate the amount of repetition of the respective step. The VR exposure was repeated for each group configuration, i.e., the within-subject factor. All logic tasks are defined in Section 3.5.2. The overall procedure is described in detail in Section 3.5.

everyone collectively, for example, starting with “Nice to have you all here. Today, you will each solve brainteasers in multiple rounds of the game”.

3.5 Experimental tasks

3.5.1 Body movement tasks

Participants engaged in three body movement tasks (i.e., moving the fingers, waving the hand, and walking in place) in front of a virtual mirror for 20 s each. The tasks were adapted from prior work (Waltemate et al., 2018; Wolf et al., 2020) aiming to familiarize participants with the virtual body and to induce an SoE by establishing visuomotor coherence (Slater et al., 2009) from both the first and third-person perspectives.

3.5.2 Logic tasks

We implemented a VR Odd-One-Out paradigm, wherein participants for each task were presented with a set of five primitive 3D objects and were required to identify the unique “odd” item within the set. We chose this type of logic task as it is a singular, non-cooperative, and non-verbal task that can be replicated in a repeated measures design by slightly changing single primitive characteristics, still requesting comparable cognitive demand for each repeated measure. Each VR exposure comprised five Odd-One-Out tasks designed with increasing difficulty to keep participants’ concentration and attention level high while avoiding boredom. Initially, participants solved three easy tasks wherein the odd object

differed in a single property, such as color, size, or geometry. Subsequently, one intermediate task challenged participants to identify an object differing in two of the named properties, and finally, one advanced task demanded participants to discern the odd object solely based on its shape, considering factors like edges or elongation (see Figure 1). For easy tasks, objects were shown for 15 s, and for intermediate and advanced tasks, objects were presented for 25 s. The tasks were created based on six different geometries (cube, cuboid, sphere, cylinder, capsule, pyramid), six distinct colors, and three sizes. Upon completion of the object presentation phase, the objects disappeared, and a tablet for solution input was displayed for a total of 20 s. Participants could click one out of five buttons, each representing one of the objects shown. We, therefore, implemented small box colliders on the self-avatars’ index distal phalanges, allowing the participants to interact with the tablet’s buttons through their virtual bodies.

3.6 Procedure

The study took place in a laboratory at the University of Würzburg and followed a controlled experimental procedure that took around 95 min ($M = 94.83$, $SD = 11.79$). All participants underwent the exact same procedure regardless of their randomly assigned test condition. This includes the creation of both a personalized and a customized self-avatar. The procedure was divided into four phases and is visualized in Figure 3.

3.6.1 The opening phase and the self-avatar creation phase

In the opening phase, participants first read the study information and gave explicit written consent for participation. In the following avatar creation phase, participants were guided to an adjacent room, where the body scan was performed as described in Section 3.4.3. Study data and body scan data were pseudonymized separately to avoid de-anonymization. Back in the main laboratory, the participants customized their stylized self-avatar as described in Section 3.4.4, while the experimenter finalized the personalized self-avatar.

3.6.2 The experiment phase

In the experiment phase, participants answered pre-questionnaires on demographic data before three consecutive VR exposures took place. For each of the three VR exposures, the experimenter demonstrated how to wear the HMD correctly and instructed participants on how to adjust the interpupillary distance to their eyes and the integrated headphones to their ears. Once immersed in the virtual environment, participants were guided via prerecorded audio instructions, following a linear procedure. Participants underwent a short vision test, and the self-avatar calibration process was conducted. Subsequently, a virtual mirror was introduced and shown in front of the participants, who then engaged in three brief body movement tasks (see Section 3.5.1). After the tasks, the mirror disappeared, and the participants were introduced to the virtual others. The virtual others (see Section 3.4.6) then appeared 2 m in front of the participants, who were then told to walk towards the group, becoming a part of it. After a detailed introduction, participants familiarized themselves with the tablet's functionality, and a set of five consecutive logic tasks were applied one by one (see Section 3.5.2). After the last logic task, the virtual others disappeared, and the participants answered seven short questions verbally in VR on spatial presence, the SoE, and self-identification. Participants then took off the HMD and answered questionnaires on the perception of the virtual others' plausibility, co-presence, and affective appraisal in the anteroom. After all VR exposures, participants answered post-questionnaires out of VR on spatial presence, the SoE, and self-identification. See Section 3.3 for an overview of all measures.

3.6.3 The closing phase

Lastly, the experimenter debriefed participants, named manipulated variables, and answered all questions.

3.7 Statistical analysis

Statistical analyses were performed with R in version 4.3.0 (R Core Team, 2023) and conducted at a significance level of $\alpha = .05$. For all models, we tested whether our control measures (Section 3.3.7) significantly impacted any dependent variables and ensured they were independent of the predictor variables. If applicable, we incorporated the respective control measures into the models described below.

We calculated mixed ANOVAs with self-avatar (between-subject) and virtual others (within-subject) as predictors and one of the dependent measures assessed after each VR exposure as the

outcome variable using the *rstatix* package (Kassambara, 2023). Effect sizes were determined using generalized eta-squared, η_G^2 (Bakeman, 2005). If the assumption of sphericity was violated, we reported Greenhouse-Geisser correction values. In cases of significant within-subject effects, we calculated pairwise *post hoc* comparisons with Bonferroni-adjusted *p*-values. For significant interaction effects, we performed pairwise *post hoc* comparisons for each between-subject factor, also using Bonferroni-adjusted *p*-values. Furthermore, we calculated Welch's *t*-tests for independent samples and Cohen's *d* effect sizes to examine the effect of self-avatar on the measures assessed once after the last VR exposure. All descriptive statistics are presented in Table 2.

4 Results

4.1 Perception of virtual others

4.1.1 Appearance and behavior plausibility (ABP)

Self-avatar, $F(1, 46) = .198$, $p = .658$, $\eta_G^2 = .002$, virtual others, $F(2, 92) = 2.60$, $p = .08$, $\eta_G^2 = .024$, and the interaction between self-avatar and virtual others, $F(2, 92) = .050$, $p = .951$, $\eta_G^2 = .000$, had no significant effect on ABP.

4.1.2 Match with virtual environment (MVE)

Virtual others had a significant main effect on MVE, $F(1.71, 78.88) = 5.198$, $p = .011$, $\eta_G^2 = .064$. On average, realistic virtual others were rated significantly higher on MVE compared to mixed virtual others, $F(47) = 3.65$, $p_{adj} = .002$, but not compared to stylized virtual others, $F(47) = 2.05$, $p_{adj} = .139$. Furthermore, self-avatar, $F(1, 46) = .036$, $p = .85$, $\eta_G^2 = .000$, and the interaction between self-avatar and virtual others, $F(1.71, 78.88) = 1.227$, $p = .294$, $\eta_G^2 = .016$, had no significant effect on MVE.

4.1.3 Co-presence

Virtual others had a significant main effect on co-presence, $F(1.72, 79.05) = 29.544$, $p < .001$, $\eta_G^2 = .142$. On average, realistic virtual others were rated significantly higher on co-presence than mixed virtual others, $F(47) = 3.06$, $p_{adj} = .011$, and mixed virtual others were rated significantly higher on co-presence compared to stylized virtual others, $F(47) = 5.57$, $p_{adj} < .001$. Furthermore, self-avatar, $F(1, 46) = 2.45$, $p = .124$, $\eta_G^2 = .038$, and the interaction between self-avatar and virtual others had no significant effect on co-presence, $F(1.72, 79.05) = 0.608$, $p = .523$, $\eta_G^2 = .003$.

4.1.4 Impression of interaction possibilities (IIP)

Virtual others had a significant main effect on IIP, $F(1.66, 76.47) = 21.973$, $p < .001$, $\eta_G^2 = .101$. On average, realistic virtual others were rated significantly higher on IIP than mixed virtual others, $F(47) = 3.37$, $p_{adj} = .005$, and mixed virtual others were rated significantly higher on IIP compared to stylized virtual others, $F(47) = 4.13$, $p_{adj} < .001$. Furthermore, self-avatar, $F(1, 46) = 0.298$, $p = .588$, $\eta_G^2 = .005$, and the interaction between self-avatar and virtual others, $F(1.66, 76.47) = 0.408$, $p = .629$, $\eta_G^2 = .002$, had no significant effect on IIP.

4.1.5 Humanness

Virtual others had a significant main effect on humanness, $F(1.66, 76.14) = 67.911$, $p < .001$, $\eta_G^2 = .331$. On average, realistic virtual others were rated significantly higher on humanness compared to mixed virtual others, $F(47) = 4.76$, $p_{adj} < .001$, and mixed virtual others were rated significantly higher compared to stylized virtual others, $F(47) = 8.62$, $p_{adj} < .001$. Furthermore, self-avatar, $F(1, 46) = .022$, $p = .884$, $\eta_G^2 < .001$, and the interaction between self-avatar and virtual others had no significant effect on humanness, $F(1.66, 76.14) = .993$, $p = .362$, $\eta_G^2 = .007$.

4.1.6 Eeriness

Self-avatar, $F(1, 46) = .661$, $p = .420$, $\eta_G^2 = .008$, virtual others, $F(1.74, 79.88) = 1.878$, $p = .165$, $\eta_G^2 = .017$, and the interaction between self-avatar and virtual others, $F(1.74, 79.88) = 1.458$, $p = .239$, $\eta_G^2 = .013$, had no significant effect on eeriness.

4.2 Perception of self-representation

4.2.1 Virtual body ownership (VBO)

Self-avatar had a significant main effect on VBO assessed in VR after each VR exposure, $F(1, 46) = 8.802$, $p = .005$, $\eta_G^2 = .139$. On average, realistic self-avatars ($M = 6.38$, $SD = 1.79$) were rated higher on VBO than stylized self-avatars ($M = 4.78$, $SD = 2.24$). Virtual others, $F(1.76, 81.01) = 0.204$, $p = .788$, $\eta_G^2 = .001$, and the interaction between self-avatar and virtual others had no significant effect on VBO in VR, $F(1.76, 81.01) = 2.646$, $p = .084$, $\eta_G^2 = .009$.

Furthermore, self-avatar had a significant effect on VBO assessed out of VR after the last VR exposure, $t(45.00) = 2.97$, $p = .005$, $d = .86$. On average, realistic self-avatars ($M = 4.68$, $SD = 1.05$) were rated higher on VBO than stylized self-avatars ($M = 3.70$, $SD = 1.23$).

4.2.2 Agency

Self-avatar, $F(1, 46) = 2.786$, $p = .102$, $\eta_G^2 = .049$, virtual others, $F(2, 92) = 0.312$, $p = .733$, $\eta_G^2 = .001$, and the interaction between self-avatar and virtual others had no significant effect on agency assessed in VR after each VR exposure, $F(2, 92) = 2.024$, $p = .138$, $\eta_G^2 = .006$.

Furthermore, self-avatar had no significant effect on agency assessed out of VR after the last VR exposure, $t(45.56) = 0.249$, $p = .804$, $d = .07$.

4.2.3 Change

Self-avatar, $F(1, 46) = 0.754$, $p = .390$, $\eta_G^2 = .016$, virtual others, $F(2, 92) = .178$, $p = .837$, $\eta_G^2 < .001$, and the interaction between self-avatar and virtual others had no significant effect on change assessed in VR after each VR exposure, $F(2, 92) = 1.737$, $p = .182$, $\eta_G^2 = .001$.

Furthermore, the covariate VR experience was significantly related to change assessed out of VR after the last VR exposure, $F(1, 45) = 8.37$, $p = .006$, $r = .40$. Self-avatar had no significant effect on change assessed out of VR after the last VR exposure, $F(1, 45) = .52$, $p = .473$, $r = .11$.

TABLE 3 Descriptive statistics for measures assessed out of VR after the last VR exposure for stylized and realistic self-avatars. All measures' ranges are [1–7].

Measures		Stylized self	Realistic self
		<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)
Sens of Embodiment	VBO	3.70 (1.23)	4.68 (1.05)
	Agency	5.75 (0.76)	5.80 (0.69)
	Change	2.97 (1.81)	3.26 (1.12)
	Self-Location	3.41 (1.08)	4.32 (0.96)
Self-Identification	Self-Identification	3.41 (1.02)	5.07 (0.93)
Spatial Presence	Spatial Presence	3.51 (0.89)	3.16 (0.74)

4.2.4 Self-location

Self-avatar and virtual others had a significant interaction effect on self-location assessed in VR after each VR exposure, $F(2, 92) = 3.145$, $p = .048$, $\eta_G^2 = .014$. For stylized self-avatars, realistic virtual others led to significantly lower ratings on self-location than stylized virtual others, $F(23) = -2.85$, $p_{adj} = .027$, and for realistic self-avatars, virtual others had no significant effect on self-location (all $p_{adj} > 0.99$). Moreover, self-avatar, $F(1, 46) = 0.225$, $p = .637$, $\eta_G^2 = .004$, and virtual others, $F(2, 92) = 1.096$, $p = .338$, $\eta_G^2 = .005$, had no significant main effect on self-location in VR.

Furthermore, self-avatar had a significant effect on self-location assessed out of VR after the last VR exposure, $t(45.37) = 3.107$, $p = .003$, $d = 0.90$. On average, realistic self-avatars ($M = 4.32$, $SD = 0.96$) were rated higher on self-location than stylized self-avatars ($M = 3.41$, $SD = 1.08$). See Table 3 for descriptive statistics.

4.2.5 Self-identification (SI)

Self-avatar had a significant main effect on SI assessed in VR after each VR exposure, $F(1, 46) = 29.149$, $p < .001$, $\eta_G^2 = .344$. On average realistic self-avatars ($M = 6.93$, $SD = 1.51$) were rated higher on SI than stylized self-avatars ($M = 4.49$, $SD = 1.90$). Also, the interaction between self-avatar and virtual others had a significant effect on SI in VR, $F(1.54, 70.66) = 6.645$, $p < .005$, $\eta_G^2 = .024$. For stylized self-avatars, stylized virtual others led to significantly higher ratings in SI compared to mixed virtual others, $F(23) = 2.99$, $p_{adj} = .02$, and realistic virtual others, $F(23) = 3.65$, $p_{adj} = .004$. For realistic self-avatars, virtual others had no significant effect on SI (all $p_{adj} > .777$).

Furthermore, self-avatar had a significant effect on SI assessed out of VR after the last VR exposure, $t(45.54) = -5.90$, $p < .001$, $d = -1.70$. On average, realistic self-avatars ($M = 5.07$, $SD = 0.93$) were rated higher on SI than stylized self-avatars ($M = 3.41$, $SD = 1.02$). See Table 3 for descriptive statistics.

4.3 Spatial presence

There were no significant effects of self-avatar, $F(1, 46) = 0.627$, $p = .432$, $\eta_G^2 = .010$, virtual others,

TABLE 4 All hypotheses and explorative evaluations as stated in Section 2.

Hypotheses		Result
Virtual Others		
H1	The manipulation of virtual others' realism and the self-avatar's realism will lead to significantly higher scores in VHP for configurations of higher congruence	Rejected
Self-Representation		
H2.1	The manipulation of the self-avatar's congruence with the participant will lead to significantly higher scores in VBO and self-identification for realistic self-avatars	Accepted
H2.2	The manipulation of the self-avatar's congruence with the participant will not significantly affect agency, change, and self-location	Partially Accepted
Spatial Presence		
H3	The manipulation of the self-avatar's realism, the virtual others' realism, and their congruence will not significantly affect spatial presence	Accepted
Explorative Evaluations		
Virtual Others		
E1.1	Evaluation of the impact of the virtual others' realism, and their congruence with the self-avatar's realism, on co-presence and the impression of interaction possibilities	
E1.2	Evaluation of the impact of the virtual others' realism, and their congruence with the self-avatar's realism, on affective appraisal and the Uncanny Valley effect	
Self-Presentation		
E2.1	Evaluation of the impact of the congruence between the self-avatar's realism and virtual other's realism on the perception of self-representation	

$F(2, 92) = 1.143$, $p = .323$, $\eta_G^2 = .006$, and the interaction between self-avatar and virtual others, $F(2, 92) = 2.118$, $p = .126$, $\eta_G^2 = .010$ on spatial presence assessed in VR after each VR exposure.

Furthermore, self-avatar had no significant effect on spatial presence assessed out of VR after the last VR exposure, $t(44.453) = -1.479$, $p = .146$, $d = -.472$.

5 Discussion

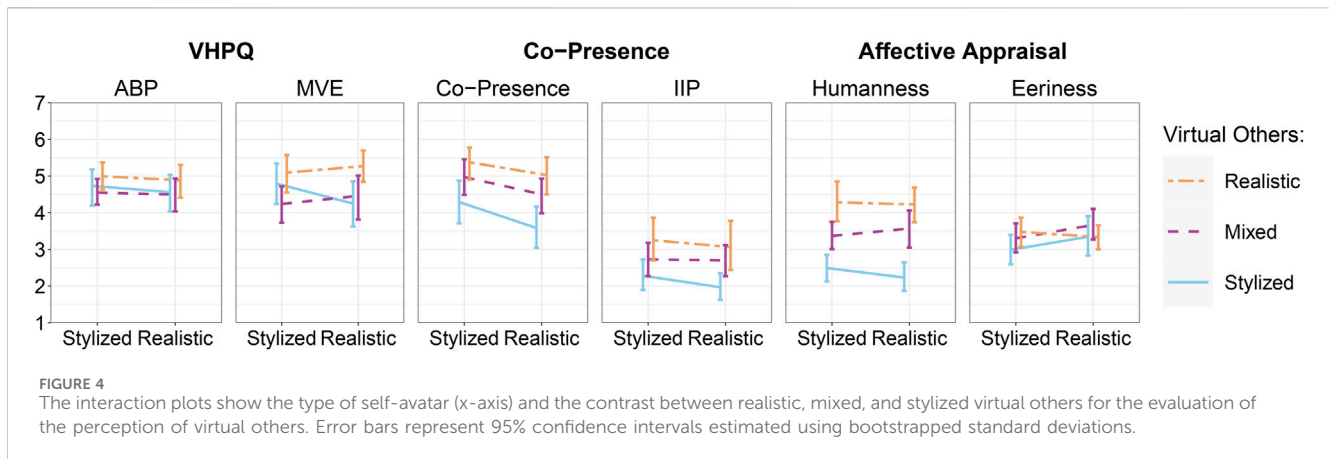
Our study investigated how the realism of self-avatars, a group of co-located virtual humans, and related (in)congruencies affected the users' VR experience regarding the perception of virtual others, self-representation, and spatial presence. In order to aid comprehension, Table 4 restates all hypotheses and explorative evaluations. The upcoming sections, will delve into our study's findings, address its limitations, and explore the implications of our research.

5.1 (In)congruencies and the perception of virtual others

5.1.1 Virtual human plausibility

Mal et al. (2022) framed the subjective feeling of how reasonable and believable a virtual human appears to a user as virtual human plausibility (VHP) arising from the congruence of its cues, eventually influencing a virtual human's appearance and behavior plausibility (ABP) and perceived match with the VE (MVE). We

hypothesized the manipulation of virtual others' realism and the self-avatar's realism to lead to significantly higher scores in VHP for configurations of higher congruence (H1). Contrary to our hypothesis, we found no differences in the perceived plausibility of appearance and behavior among virtual others between realistic, mixed, and stylized configurations, homogeneous and mixed groups, or conditions with (in)congruent realism between self-avatars and virtual others. Despite our deliberate manipulation of realism introducing higher order cognitive incongruencies (Latoschik and Wienrich, 2022), as described in Section 2.2.2, these did not seem to affect participants' subjective perceptions of how reasonable and believable the appearance and behavior of the group of others appeared to them. Participants demonstrated notable flexibility in accepting varying styles as plausible for a group of virtual humans. Participants may have accommodated incongruencies on a cognitive level based on their habituation to virtual environments simulating alternative realities. Since we did not specify whether the virtual others were real humans or agents (see Section 3.4.6.3), the rule-based behavior might have suggested that the virtual others were computer-controlled. Consequently, perceiving the virtual others as agents may have reduced the need for them to have a realistic, human-like appearance to be considered believable in an alternative, virtual reality. However, this is an interpretative approach that suggests the need for further research. We also consider that the preliminary VHP questionnaire may not have been sensitive to our manipulation, as subsequent findings clearly show a significant manipulation of how virtual others were perceived based on other qualia.



In terms of the virtual others' match with the virtual environment (MVE), a homogeneous group of realistic virtual others received the highest ratings for MVE and was rated significantly higher than the mixed configuration of virtual others, but not the homogeneous stylized configuration. This overall suggests that the virtual environment better aligned with the appearance of realistic virtual humans. Further analyzing the interaction plot presented in Figure 4 for MVE indicates that for realistic self-avatars, the (in)congruent stylized configuration of virtual others received the lowest ratings. Conversely, with stylized self-avatars, the congruent stylized configuration of virtual others was more favorably accepted and rated as fitting better within the virtual environment. However, we approach this interpretation cautiously as the interaction between self-avatar and virtual others did not yield significant differences. We suggest future research to delve deeper into the descriptively indicated impact of self-representation on the evaluation of virtual others' match with a virtual environment.

5.1.2 Co-presence and impression of interaction possibilities (IIP)

In an exploratory evaluation, we investigated the impact of the virtual others' realism, and their congruence with the self-avatar's realism, on co-presence and the impression of interaction possibilities (E1.1). The realism of virtual others had a notable effect on co-presence. Our results indicate that configurations featuring realistic virtual others greatly enhanced the perceived co-presence with the group. Specifically, group configurations with exclusively realistic others received higher ratings than the mixed others, and mixed others were rated higher than the all-stylized configuration. This effect was consistent for both self-avatars, suggesting that the (in)congruence between the realism of the self-avatar and the group of others did not have an impact on the sense of "being there together" (Schroeder, 2002). While previous research on the impact of the appearance of virtual humans on co-presence has yielded heterogeneous results in general, our findings align closely with the work of Zibrek et al. (2017). The authors discovered that participants preferred realistic virtual humans over stylized representations. While the named work utilized a single virtual human as an agent, we can confirm these findings in a group setting. Further aligning with Zibrek et al. (2017), our

exploratory evaluation of eeriness did not suggest increased feelings of unappealing for any of the styles, which might be the base for realism in virtual humans to benefit co-presence. Further, the systematic review of Oh et al. (2018) indicated the congruence between realism in appearance and behavior to positively predict co-presence. In our work, the virtual others displayed rather realistic body movements either exported from Mixamo (Adobe Systems Inc., 2023) or captured with a state-of-the-art motion capture system (Movella Inc., 2022). These movements might have been more congruent with the realistic others than the abstract ones, eventually leading to increased co-presence.

Interestingly, participants also reported that their impression of interaction possibilities with virtual others was higher for configurations featuring realistic virtual others, an effect aligning closely with the co-presence trend. Notably, the overall ratings for IIP were rather low, possibly due to the rule-based behavior of the virtual others and the explicit framing that verbal communication with the others was not allowed (see Section 3.4.6), both limiting the interaction possibilities. Nonetheless, participants still perceived that realistic virtual humans offered more contingencies for interaction. We assume the impression of co-presence and interaction possibility with virtual others to be affected by their congruence with real-world knowledge (i.e., being realistic), aligning with the impression that (real) humans are capable of engaging in interactions (interaction possibilities) while being in a place together (co-presence).

However, our findings contrast the work of Latoschik et al. (2019), who suggested potential incongruencies with participants' expectations caused them to focus more intensely on the surrounding agents when they had mixed appearances. The key distinction between the cited study and ours is that Latoschik et al. (2019) utilized a passive ambient crowd, representing an environmental surrounding rather than virtual entities with which one might want to interact directly. From these differences, we can derive that to enhance users' interest in an SVE, providing mixed avatar appearances in an ambient crowd might be beneficial (Latoschik et al., 2019). In group situations, on the other hand, where the virtual others are in a realm that might enable direct interaction, even if it was framed as not allowed, a realistic appearance may increase the impression of interaction possibilities and, likewise, co-presence.

5.1.3 Affective appraisal

We exploratively evaluated the impact of virtual others' realism, and their congruence with the self-avatar's realism on affective appraisal and the Uncanny Valley effect (E1.2). Therefore, we investigated whether groups of virtual others, comprising varying numbers of realistic or stylized virtual humans, exhibited different levels of humanness and whether they appeared eerie to the users. Supporting the effectiveness of our manipulation, virtual others were perceived as more human when they included a higher proportion of realistic virtual humans in the group's configuration. Further, we did not observe any differences in the perceived eeriness attributed to the groups' configurations for both self-avatar types. We deduce that we did not encounter an uncanny valley effect for the virtual others and assume that the perceived eeriness did not interfere with our overall results. Simultaneously, there is no evidence to suggest that incongruent styles between the self-avatar and the virtual others led to an eerie perception of virtual others.

5.2 (In)congruencies and the perception of the self-representation

5.2.1 The Self-Avatar's congruence with the participant

Concerning the self-avatar's congruence with the participant, we hypothesized the manipulation of the self-avatar's congruence with the participant to lead to significantly higher scores in VBO and self-identification for realistic self-avatars (H2.1), while we expected no significant differences in agency, change and self-location (H2.2). In line with H2.1, realistic self-avatars resulted in a significant increase in self-identification and VBO, measured both in VR and after all VR exposures (post-VR). We attribute these differences between the perception of realistic and stylized self-avatars to the influence of two higher-order cues contributing to the cognitive processing of the self-avatar's congruence with the participant. First, consistent with prior research suggesting that avatar realism enhances self-identification and VBO (Latoschik et al., 2017; Salagean et al., 2023), a (more) realistic self-avatar may have provided visual cues that created the impression of sufficient human likeness congruent with real-world knowledge on human anatomy and composition (Lugrin et al., 2015). Secondly, our results align with previous studies that have highlighted personalization as a key factor influencing VBO and self-identification (Waltemate et al., 2018; Salagean et al., 2023). We assume that the scan-based personalization process employed for realistic self-avatars offered a higher degree of truthfulness compared to the customization process used for stylized self-avatars. Eventually, participants perceived a greater similarity between themselves and the realistic self-avatars and potentially attributed more personal characteristics to them (Wolf et al., 2022a). Interestingly, in contradiction to H2.2, realistic self-avatars also resulted in increased self-location, an effect we measured after the last VR exposure. We consider the named visual cues, to likely also have enhanced the spatial experience of being inside a virtual body. However, we did not observe the same effect in the repeated in-VR measurements, limiting the significance of the finding and leading us to propose further investigation. Overall, we assume increased realism and truthfulness have resulted in visual cues congruent with the participants' real-world

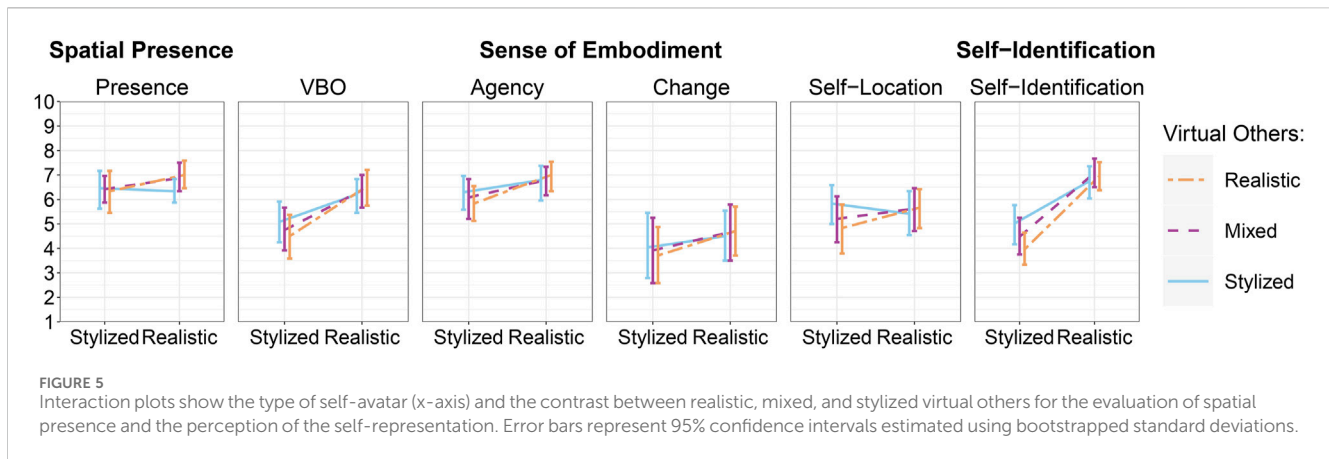
experiences concerning their own physical body, presuming one can identify with the virtual body, enhancing the sense of being inside and having it as one's own.

Furthermore, in line with H2.2, the self-avatar's congruence with the participant did not yield differences in agency and change. Concerning agency, we posit that our technical system provided comparable embodiment configurations for both self-avatar types and, consequently, the same contingencies for controlling the virtual body during active movement (Kilteni et al., 2012). We suggest the sense of agency over the virtual body primarily arose from the congruence of visuomotor and proprioceptive cues and, therefore, was less susceptible to manipulation from top-down influences. These results align with the recent meta-analysis by Mottelson et al. (2023), indicating limited effects of the avatar's appearance on agency. In terms of change, our findings indicated overall low values, which might be attributed to our efforts to minimize disparities between participants' actual body appearances and their self-avatars through the process of individualization, eventually leading to negligible alterations in the overall body schema.

5.2.2 The congruence between the self-avatar and virtual others

An exploratory evaluation of the congruence between self-avatars' realism and the virtual others' realism revealed notable effects on the perception of self-representation (E2.1). Interestingly, only for those participants embodying a stylized self-avatar, being together with an incongruent group of realistic virtual others led to lower ratings of self-location. A comparable pattern occurred for self-identification as both configurations of virtual others containing incongruent realistic virtual humans, i.e., mixed and realistic, seemed to hinder the process of identifying with the stylized self-avatar. Assuming self-location as well as self-identification to be concerned with the relationship between one's self and one's body (Kilteni et al., 2012), it is noteworthy that higher-order (in) congruent visual cues that are not within the ego-central referential frame of one's (virtual) body, can have an effect on self-location and identification. The presence of realistic virtual others in the virtual environment may have accentuated the contrast between visual cues in rendering realistic virtual humans (others) and the stylized self-representation (self-avatar) not resembling realistic geometry and textures. This increased awareness may have further highlighted the incongruence between the participants' real-world bodies, as they are realistic *per se*, and the embodied stylized digital self-representation. The interaction plot (Figure 5) also indicates a similar tendency for VBO; however, the interaction was not statistically significant ($p = .084$). We are not aware of a comparable finding that the congruence between the realism of self-avatars and virtual others significantly impacts VBO, self-location, or identification with self-representation. It is critical to note that the questions used for self-location and self-identification (Fiedler et al., 2023) have not yet been validated. Therefore, our findings necessitate additional investigation to draw further conclusions.

In contrast to marginal indications from Latoschik et al. (2017), we did not reveal an impact of virtual others' realism on participants' change in the self-perceived body schema. Besides differences in the executed tasks and the mirror exposure time between the named



study and our experimental procedure, this discrepancy may be attributed to the individualized nature of our self-avatars, sharing fewer dissimilarities with participants' physical bodies, as opposed to the generalized avatars used in the work of [Latoschik et al. \(2017\)](#). This individualization may have resulted in fewer alterations to the overall body schema, potentially diminishing or eliminating the influence of others and their congruence with the self-avatar on the change measure.

5.3 (In)congruencies and spatial presence

In line with H3, the realism of the avatar did not lead to differences in spatial presence, nor did the style configuration of virtual others and their congruence. First, this finding aligns with the prevailing conceptualization of spatial presence as being primarily driven by bottom-up factors, influenced by the system's immersion ([Slater and Wilbur, 1997](#); [Slater et al., 2022](#)), a factor we consistently maintained for all conditions. Secondly, in relation to the CaP model, our results demonstrate comparable rendering of congruent spatial cues across all avatars and configurations of virtual others. The manipulation of higher-order visual cues pertaining to the realism and individualization of virtual humans appeared to predominantly influence the plausibility of one's virtual body as well as the group of virtual others. However, these manipulations contributed less to a general sense of "being there" within the virtual environment. The recent experiment by [Salagean et al. \(2023\)](#) supports our observations for the self-avatar, as the authors reported personalization, realism, or their interaction in self-avatars not to affect presence. Interestingly, we provide evidence for a distinction between self-location, referring to the relationship between one's self and one's body ([Kilteni et al., 2012](#)) and spatial presences, concerned with the relationship between one's self and the environment ([Wirth et al., 2007](#)). Our results indicate that manipulating the visual cues of the self-avatar and its congruence with virtual others impacted one's spatial experience of being inside a virtual body rather than being inside a virtual environment. However, it is worth noting that there is conflicting evidence in the literature, with some studies suggesting that spatial presence may indeed benefit from realism in avatars ([Weidner et al., 2023](#)) as well as from personalization ([Waltemate et al., 2018](#)). We recommend future work to delve deeper into exploring the relationship between

the sense of being there in the virtual environment, the appearance of virtual humans, and the resulting congruence of spatial cues. This investigation might not only list related studies but also compare effect sizes and system configurations, which could serve as a valuable extension to the review of [Weidner et al. \(2023\)](#).

5.4 The impact of realism in a group of virtual humans

A substantial body of research has been concerned with understanding how the realism of avatars and agents impacts users' virtual experiences and their evaluation of co-located virtual others ([Nowak and Fox, 2018](#); [Weidner et al., 2023](#)). Yet, there has not been further investigation into the effects of being co-located with multiple virtual humans of different styles in an avatar-mediated VR setting, considering both the realism of co-located virtual humans, the realism of the self-avatar, and their (in)congruencies. Our results emphasize the benefits of a realistic appearance for a group of virtual humans. First, for virtual others, we identified that the group's overall realism, i.e., its congruence with participants' real-life experiences, benefits the VR experience in terms of co-presence and perceived interaction possibilities, while the internal congruence of the group's style configuration did not yield significant differences. Second, a realistic self-avatar congruent with the participants' real-world experiences concerning their own physical bodies was important for identifying with and accepting the virtual body as one's own. Lastly, only for stylized self-avatars did incongruencies between the self-avatar and the virtual others' configuration have an adverse effect on the relationship between one's self and body. This might even indicate that realistic self-avatars can help to adapt to incongruencies between the realism of self and others. We thus infer that the further development and increasing accessibility of technologies to provide lifelike avatars, e.g., through 3D reconstruction, can help to improve the VR experience in virtual group settings.

5.5 Limitations and future work

In order to evaluate the implications of our results, we will outline certain limitations of our study and use them as a basis for future research. First, in an approach to ensure high experimental

control and consistency of behavior across all configurations of virtual others, we decided to limit the interactivity between participants and the virtual others. Therefore, we implemented a rule-based, event-driven approach to generate virtual others' behavior (see Section 3.4.6), not allowing for facial expressions and direct interaction. While the work of (Volonte et al., 2018) indicates comparable visual attention between conversational and non-conversational animations in virtual humans, our approach may have limited the ecological validity (Baumeister and Vohs, 2023) of the agents' displayed behavior, and may have conveyed false affordances, i.e., "possibilities for action and interactions that seem achievable but that are actually not carried out by the simulation" (Dufresne et al., 2024, p. 3). This limitation was also reflected in low values on the impression of interaction possibilities. We suggest future work to deepen insight by utilizing different styles of virtual others in an extended scenario with social interaction, e.g., by modeling robust virtual human conversations (Rossen et al., 2009), enabling AI-controlled virtual humans (Wienrich and Latoschik, 2021), or implement a real social interaction between multiple users of an SVE.

Second, we provided different types of individualization for the types of virtual humans to shape the manipulation of the self-avatar's congruence with the participant's physical body. We assume that the personalization process resulted in a higher level of truthfulness compared to the customization process, as the personalization objectively resembled more details of the participant's actual appearance than customization. However, the two methods also differ in how participants have agency over their avatar's creation. While the personalization process primarily aims to mirror the current state of the participant's appearance, customization empowers participants to create their avatars based on their own goals, which may or may not prioritize a resemblance to reality. To maintain a reasonable level of consistency, we explicitly instructed the participants to choose configurations similar to their physical appearance. Future work should carefully consider the impact of differences in agency over the individualization process on an individual's VR experience.

Lastly, given a sample size of $N = 48$, our analyses may not have had the intended power to reveal all hypothesized effects. A *post hoc* sensitivity analysis ($\alpha = 0.05$, $1 - \beta = 0.8$, $N = 48$) using G*Power version 3.1.9.6 (Faul et al., 2007) revealed that a two-grouped (self-avatar) mixed ANOVA with three repeated measures (virtual others) would be sensitive to medium to large effects for between-group effects ($f = 0.34$) and small-to-medium effects for within-group and within-between-group interaction effects ($f = 0.19$) (Cohen, 2013). Likewise, for the measures after all VR exposures, a *post hoc* sensitivity analysis for a *t*-test with independent samples ($\alpha = 0.05$, $1 - \beta = 0.8$, $N_1 = 24$, $N_2 = 24$) would be sensitive to medium to large effects ($d = 0.73$). We cannot rule out a type 2 error, especially for small to medium between-group effects.

5.6 Implications

Our findings carry important implications for the design of social virtual environments (SVEs) incorporating virtual humans of varying degrees of realism and anthropomorphism. Overall,

participants showed remarkable flexibility in accepting varying styles in a mixed group of virtual humans and its impact on the VR experience. We assumed a homogeneous group configuration to be more congruent than a mixed one, an (in)congruence that can be accessed by directly comparing others within the VE. However, these aspects of virtual others, namely, their homogeneity of group configuration, did not yield differences in our qualia space, offering enormous versatility for the design of shared virtual applications. We infer that mixing styles in a group of virtual others does not *per se* lead to a less plausible user experience. However, another congruence in virtual others proved important; their lifelikeness or realism, respectively. Following our assumption, realistic virtual others may be congruent with the participant's real-world experiences and expectations, elevating co-presence and the impression of interaction possibilities. We propose considering realistic appearances for virtual others to enhance the realism and effectiveness of interactions (Kyriltsias and Michael-Grigoriou, 2022), e.g., in collaborative or educational workspaces, eventually fostering a more natural and productive immersive social interaction (Scavarelli et al., 2021; Orel, 2022). Furthermore, for self-representation, applications can benefit from congruence with the participant's real-world experiences. Within the referential frame of one's physical body, a realistic and truthful self-avatar can enhance users' sense of owning and identification with the virtual body, which, e.g., might be particularly relevant in the therapeutic context supporting patients in developing realistic self-perception and a positive body image (Döllinger et al., 2019). However, relating to certain use cases, also avatars incongruent or "dissimilar" with the user's physical appearance could help to solve incongruencies within other referential frames, e.g., by establishing consistency between the self-avatar and other elements of the VE, or the overall framing or habituation of the experience (Cheymol et al., 2023). Lastly, the incongruencies between the self-avatar of stylized appearance and realistic virtual others could lead to an altered self-perception with potential negative impacts on the relationship between our self and body. Considering this effect in the development of SVEs, future work might aim at locally transforming the style of all virtual others to a stylized appearance only for those participants embodying a stylized avatar. However, implementing such a transformation without the participant's knowledge can raise ethical concerns, and participants should be actively informed about the potential modifications to their virtual environment.

5.7 Conclusion

Virtual humans play a pivotal role in a wide range of social virtual environments, and a remarkable body of research on the visualization of avatars and agents highlights the impact of virtual humans' visual cues contributing to a user's individual VR experience. Our work aligns with current theories and models, emphasizing the role of congruence and plausibility in influencing users' virtual experiences and effects, by focusing on the intricate dynamics emerging from (in)congruent styles of a group of virtual humans, including multiple co-located others

(agents), and one's digital self-representation (avatar). We indicate groups of virtual others of higher realism to increase the feeling of co-presence and the impression of interaction possibilities, while (in)congruencies in the homogeneity of the group did not cause considerable effects. Furthermore, realistic self-avatars congruent with participants' own physical bodies yielded notable benefits for virtual body ownership and self-identification with the digital representation. Notably, the incongruence between a stylized self-avatar and a group of realistic virtual others resulted in diminished ratings of self-location and self-identification, suggesting an adverse effect on the relationship between one's self and body. In conclusion, a group of virtual humans varying in realism can result in a multitude of (in)congruent visual cues impacting a VR experience. Therefore, considering these (in)congruencies to tailor a virtual experience to the application's purpose and target audience constitutes an important yet challenging task to which we contribute empirical evidence and the discussion of their implications.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics Committee of the Institute of Human-Computer-Media at the Faculty of Human Sciences of the Julius Maximilian University of Würzburg. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

David Mal: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Visualization, Writing—original draft, Software, Writing—review and editing. Nina Döllinger: Formal Analysis, Writing—review and editing. Erik Wolf: Methodology, Software, Writing—review and editing. Stephan Wenninger: Methodology, Software, Writing—review and editing. Mario Botsch: Methodology, Software, Writing—review and editing.

References

- Achenbach, J., Waltemate, T., Latoschik, M. E., and Botsch, M. (2017). "Fast generation of realistic virtual humans," in *Proceedings of the 23rd ACM symposium on virtual reality software and technology* (New York, NY, USA: Association for Computing Machinery), 1–10. doi:10.1145/3139131.3139154
- Adobe Systems Inc (2023). *Mixamo*. Available at: <https://www.mixamo.com/> (Accessed November, 2023).
- Agisoft (2021). *Metashape pro*. Available at: <http://www.agisoft.com> (Accessed November, 2023).
- Bailenson, J. N., and Blascovich, J. (2004). "Avatars," in *Encyclopedia of human-computer interaction* (Great Barrington, MA, USA: Berkshire Publishing Group), 64–68.
- Bailenson, J. N., Swinth, K., Hoyt, C., Persky, S., Dimov, A., and Blascovich, J. (2005). The independent and interactive effects of embodied-agent appearance and

Carolin Wienrich: Funding acquisition, Project administration, Resources, Supervision, Writing—review and editing. Marc Erich Latoschik: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing—review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research has been funded by the German Federal Ministry of Education and Research (BMBF) in the projects ViTraS (Grant 16SV8219 and 16SV8225) and ViLeArn More (Grant 16DHB2214), by the German Federal Ministry of Labour and Social Affairs (BMAS) in the project AIL AT WORK (Grant DK1.00.00030.21), and by the Bavarian State Ministry For Digital Affairs in the project XR Hub (Grant A5-3822-2-16). It was further supported by the Open Access Publication Fund of the University of Würzburg. Erik Wolf gratefully acknowledges a Meta Research Ph.D. Fellowship.

Acknowledgments

We thank Lena Holderrieth, Florian Kern, and Jonathan Tschanter, who are with the Human-Computer Interaction Group of the University of Würzburg, for their help in developing the Unity application and data collection. A preprint of this work has been published on 11 March 2024 (Mal et al., 2024).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of *Frontiers*, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence* 14, 379–393. doi:10.1162/105474605774785235
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behav. Res. Methods* 37, 379–384. doi:10.3758/BF03192707
- Baumeister, R., and Vohs, K. D. (2023). *Encyclopedia of social psychology*. Editor R. F. Baumeister and K. D. Vohs (Thousand Oaks, CA: SAGE Publications, Inc.), 276. doi:10.4135/9781412956253
- Bente, G., Rüggenberg, S., Krämer, N. C., and Eschenburg, F. (2008). Avatar-mediated networking: increasing social presence and interpersonal trust in net-based collaborations. *Hum. Commun. Res.* 34, 287–318. doi:10.1111/j.1468-2958.2008.00322.x
- Bergström, I., Azevedo, S., Papiotis, P., Saldanha, N., and Slater, M. (2017). The plausibility of a string quartet performance in virtual reality. *IEEE Trans. Vis. Comput. Graph.* 23, 1352–1359. doi:10.1109/TVCG.2017.2657138
- Blender Foundation (2019). Blender - an open source 3D creation suite. Available at: <https://www.blender.org/> (Accessed November, 2023).
- Bouaziz, S., Tagliasacchi, A., and Pauly, M. (2014). “Dynamic 2D/3D registration,” in *Eurographics tutorials*, 1–17. doi:10.1145/2504435.2504456
- Bouchard, S., Robillard, G., St-Jacques, J., Dumoulin, S., Patry, M.-J., and Renaud, P. (2004). “Reliability and validity of a single-item measure of presence in vr,” in *The 3rd IEEE international workshop on haptic, audio and visual environments and their applications* (IEEE), 59–61. doi:10.1109/HAVE.2004.1391882
- Boykov, Y., Veksler, O., and Zabih, R. (2001). Fast approximate energy minimization via graph cuts. *ACM Trans. Pattern Analysis Mach. Intell.* 23, 1222–1239. doi:10.1109/34.969114
- Burden, D., and Savin-Baden, M. (2019). *Virtual humans: today and tomorrow*. New York, NY: CRC Press.
- Butz, M., Hepperle, D., and Wölfel, M. (2022). “Influence of visual appearance of agents on presence, attractiveness, and agency in virtual reality,” in *ArtsIT, interactivity and game creation* (Springer International Publishing), 44–60.
- Cheymol, A., Fribourg, R., Lécuyer, A., Normand, J.-M., and Argelaguet, F. (2023). Beyond my real body: characterization, impacts, applications and perspectives of “dissimilar” avatars in virtual reality. *IEEE Trans. Vis. Comput. Graph.* 29, 4426–4437. doi:10.1109/TVCG.2023.3320209
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. New York, NY: Routledge. doi:10.4324/9780203771587
- Debarba, H. G., Chagué, S., and Charbonnier, C. (2022). On the plausibility of virtual body animation features in virtual reality. *IEEE Trans. Vis. Comput. Graph.* 28, 1880–1893. doi:10.1109/TVCG.2020.3025175
- Doerner, R., Broll, W., Grimm, P., and Jung, B. (2022). *Virtual and augmented reality (VR/AR): foundations and methods of extended realities (XR)* (Springer International Publishing). doi:10.1007/978-3-030-79062-2
- Döllinger, N., Beck, M., Wolf, E., Mal, D., Botsch, M., Latoschik, M. E., et al. (2023). “‘If it’s not me it doesn’t make a difference’ – the impact of avatar personalization on user experience and body awareness in virtual reality,” in *2023 IEEE international symposium on mixed and augmented reality (ISMAR)*, 483–492. doi:10.1109/ISMAR59233.2023.00063
- Döllinger, N., Wienrich, C., Wolf, E., Botsch, M., and Latoschik, M. E. (2019). “ViTraS - virtual reality therapy by stimulation of modulated body image - project outline,” in *Mensch und computer 2019 - workshopband, hamburg, germany*. doi:10.18420/muc2019-ws-633
- Ducheneaut, N., Wen, M.-H., Yee, N., and Wadley, G. (2009). “Body and mind: a study of avatar personalization in three virtual worlds,” in *Proceedings of the SIGCHI conference on human factors in computing systems* (New York, NY, USA: Association for Computing Machinery), 1151–1160. CHI '09. doi:10.1145/1518701.1518877
- Dufresne, F., Dubosc, C., Gorisse, G., and Christmann, O. (2024). “Understanding the impact of coherence between virtual representations and possible interactions on embodiment in vr: an affordance perspective,” in *Extended abstracts of the 2024 CHI conference on human factors in computing systems* (New York, NY: Association for Computing Machinery). CHI EA '24. doi:10.1145/3613905.3650752
- Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). *G* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences*. Springer. doi:10.3758/BF03193146
- Fiedler, M. L., Wolf, E., Döllinger, N., Botsch, M., Latoschik, M. E., and Wienrich, C. (2023). “Embodiment and personalization for self-identification with virtual humans,” in *2023 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*, 799–800. doi:10.1109/VRW58643.2023.00242
- Freeman, G., and Maloney, D. (2021). “Body, avatar, and me: the presentation and perception of self in social virtual reality,” in *Proc. ACM on human-computer interaction* (New York, NY, USA: Association for Computing Machinery), 239:1–239:27. doi:10.1145/3432938
- Fribourg, R., Argelaguet, F., Lécuyer, A., and Hoyet, L. (2020). Avatar and sense of embodiment: studying the relative preference between appearance, control and point of view. *IEEE Trans. Vis. Comput. Graph.* 26, 2062–2072. doi:10.1109/TVCG.2020.2973077
- Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A., and Sasse, M. A. (2003). “The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment,” in *Proceedings of the conference on Human factors in computing systems - CHI '03* (New York, NY: ACM Press), 529. doi:10.1145/642611.642703
- González-Franco, M., Steed, A., Hoogendyk, S., and Ofek, E. (2020). Using facial animation to increase the enforcement and avatar self-identification. *IEEE Trans. Vis. Comput. Graph.* 26, 2023–2029. doi:10.1109/tvcg.2020.2973075
- Gorisse, G., Christmann, O., Houzangbe, S., and Richir, S. (2019). From robot to virtual doppelgänger: impact of visual fidelity of avatars controlled in third-person perspective on embodiment and behavior in immersive virtual environments. *Front. Robotics AI* 6, 8. doi:10.3389/frobt.2019.00008
- Hepperle, D., Purps, C. F., Deuchler, J., and Wölfel, M. (2022). Aspects of visual avatar appearance: self-representation, display type, and uncanny valley. *Vis. Comput.* 38, 1227–1244. doi:10.1007/s00371-021-02151-0
- Ho, C.-C., and MacDorman, K. F. (2017). Measuring the Uncanny Valley Effect. *Int. J. Soc. Robotics* 9, 129–139. doi:10.1007/s12369-016-0380-9
- Kassambara, A. (2023). Rstatix: pipe-friendly framework for basic statistical tests. *R package version 0.7.2*. Available at: <https://CRAN.R-project.org/package=rstatix/> (Accessed November, 2023).
- Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *Int. J. Aviat. Psychol.* 3, 203–220. doi:10.1207/s15327108ijap0303_3
- Kiltner, K., Groten, R., and Slater, M. (2012). The sense of embodiment in virtual reality. *Presence Teleoperators Virtual Environ.* 21, 373–387. doi:10.1162/PRES_a_00124
- Kim, K., Maloney, D., Bruder, G., Bailenson, J. N., and Welch, G. F. (2017). The effects of virtual human’s spatial and behavioral coherence with physical objects on social presence in AR. *Comput. Animat. Virtual Worlds* 28, e1771. doi:10.1002/cav.1771
- Kolesnichenko, A., McVeigh-Schultz, J., and Isbister, K. (2019). “Understanding emerging design practices for avatar systems in the commercial social VR ecology,” in *Proceedings of the 2019 on designing interactive systems conference* (New York, NY, USA: Association for Computing Machinery), 241–252. doi:10.1145/3322276.3322352
- Kullmann, P., Menzel, T., Botsch, M., and Latoschik, M. E. (2023). “An evaluation of other-avatar facial animation methods for social VR,” in *Extended abstracts of the 2023 CHI conference on human factors in computing systems*. doi:10.1145/3544549.3585617
- Kyrlitsias, C., and Michael-Grigoriou, D. (2022). Social interaction with agents and avatars in immersive virtual environments: a survey. *Front. Virtual Real.* 2. doi:10.3389/frvir.2021.786665
- Latoschik, M. E., Kern, F., Stauffert, J.-P., Bartl, A., Botsch, M., and Lugrin, J.-L. (2019). Not alone here?! Scalability and user experience of embodied ambient crowds in distributed social virtual reality. *IEEE Trans. Vis. Comput. Graph.* 25, 2134–2144. doi:10.1109/TVCG.2019.2899250
- Latoschik, M. E., Roth, D., Gall, D., Achenbach, J., Waltemate, T., and Botsch, M. (2017). “The effect of avatar realism in immersive social virtual realities,” in *Proceedings of the 23rd ACM symposium on virtual reality software and technology* (Gothenburg, Sweden: Association for Computing Machinery), 1–10. doi:10.1145/3139131.3139156
- Latoschik, M. E., and Wienrich, C. (2022). Congruence and plausibility, not presence?! Pivotal conditions for XR experiences and effects, a novel approach. *Front. Virtual Real.* 3. doi:10.3389/frvir.2022.694433
- Lee, H.-W., Chang, K., Uhm, J.-P., and Owiro, E. (2023). How avatar identification affects enjoyment in the metaverse: the roles of avatar customization and social engagement. *Cyberpsychology, Behav. Soc. Netw.* 26, 255–262. doi:10.1089/cyber.2022.0257
- LimeSurvey GmbH (2020). *LimeSurvey: an open source survey tool*. Available at: <https://www.limesurvey.org/Version4.5> (Accessed November, 2023).
- Liu, Q., and Steed, A. (2021). Social virtual reality platform comparison and evaluation using a guided group walkthrough method. *Front. Virtual Real.* 2. doi:10.3389/frvir.2021.668181
- Longo, M. R., Schüür, F., Kammers, M. P., Tsakiris, M., and Haggard, P. (2008). What is embodiment? a psychometric approach. *Cognition* 107, 978–998. doi:10.1016/j.cognition.2007.12.004
- Lugrin, J.-L., Latt, J., and Latoschik, M. E. (2015). “Anthropomorphism and illusion of virtual body ownership,” in *Proceedings of the 25th international conference on artificial reality and telexistence and 20th eurographics symposium on virtual environments* (Kyoto, Japan: Eurographics Association), 1–8.
- Mal, D. (2020). “The impact of social interactions on an embodied individual’s self-perception in virtual environments,” in *2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*, 545–546. doi:10.1109/VRW50115.2020.00124

- Mal, D., Döllinger, N., Wolf, E., Wenninger, S., Botsch, M., Wienrich, C., et al. (2024). Am I the odd one? Exploring (in)congruencies in the realism of avatars and virtual others in virtual reality. *arXiv*. doi:10.48550/arXiv.2403.07122
- Mal, D., Wolf, E., Döllinger, N., Botsch, M., Wienrich, C., and Latoschik, M. E. (2022). "Virtual human coherence and plausibility – towards a validated scale," in *2022 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*, 788–789. doi:10.1109/VRW55335.2022.00245
- Mal, D., Wolf, E., Döllinger, N., Wienrich, C., and Latoschik, M. E. (2023). The impact of avatar and environment congruence on plausibility, embodiment, presence, and the proteus effect in virtual reality. *IEEE Trans. Vis. Comput. Graph.* 29, 2358–2368. doi:10.1109/TVCG.2023.3247089
- Maselli, A., and Slater, M. (2013). The building blocks of the full body ownership illusion. *Front. Hum. Neurosci.* 7, 83. doi:10.3389/fnhum.2013.00083
- McDonnell, R., Breidt, M., and Bühlhoff, H. H. (2012). Render me real? Investigating the effect of render style on the perception of animated virtual humans. *ACM Trans. Graph.* 31 (91), 1–11. doi:10.1145/2185520.2185587
- McVeigh-Schultz, J., Kolesnichenko, A., and Isbister, K. (2019). "Shaping pro-social interaction in VR: an emerging design Framework," in *Proceedings of the 2019 CHI conference on human factors in computing systems* (Glasgow Scotland UK: ACM), 1–12. doi:10.1145/3290605.3300794
- Monk, E. (2019). Monk skin tone scale. Available at: <https://skintone.google> (Accessed November, 2023).
- Mori, M., MacDorman, K. F., and Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics Automation Mag.* 19, 98–100. doi:10.1109/MRA.2012.2192811
- Mottelson, A., Muresan, A., Hornbæk, K., and Makransky, G. (2023). A systematic review and meta-analysis of the effectiveness of body ownership illusions in virtual reality. *ACM Trans. Comput.-Hum. Interact.* 30, 1–42. doi:10.1145/3590767
- Movella Inc (2022). *Xsens MVN Record*. Available at: <https://www.xsens.com/> (Accessed November, 2023).
- Nowak, K. L., and Fox, J. (2018). Avatars and computer-mediated communication: a review of the definitions, uses, and effects of digital representations on communication. *Rev. Commun. Res.* 6, 30–53. doi:10.12840/issn.2255-4165.2018.06.01.015
- Oh, C. S., Bailenson, J. N., and Welch, G. F. (2018). A systematic review of social presence: definition, antecedents, and implications. *Front. Robotics AI* 5, 114. doi:10.3389/frbot.2018.00114
- Orel, M. (2022). "The virtual reality workplace," in *Collaboration potential in virtual reality (VR) office space: transforming the workplace of tomorrow* (Cham: Springer International Publishing), 35–73. doi:10.1007/978-3-031-08180-4_4
- Peck, T. C., and Gonzalez-Franco, M. (2021). Avatar embodiment. a standardized questionnaire. *Front. Virtual Real.* 1. doi:10.3389/frvir.2020.575943
- Pérez, P., Gangnet, M., and Blake, A. (2003). Poisson image editing. *ACM Trans. Comput. Graph.* 22, 313–318. doi:10.1145/1201775.882269
- Poeschl, S., and Doering, N. (2015). Measuring co-presence and social presence in virtual environments – psychometric construction of a German scale for a fear of public speaking scenario. *Annu. Rev. Cybertherapy Telemedicine* 2015, 58–63. doi:10.3233/978-1-61499-595-1-58
- Ratan, R., Beyea, D., Li, B. J., and Graciano, L. (2019). Avatar characteristics induce users' behavioral conformity with small-to-medium effect sizes: a meta-analysis of the Proteus effect. *Media Psychol.* 23, 651–675. doi:10.1080/15213269.2019.1623698
- R Core Team (2023). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: <https://www.r-project.org/> (Accessed November, 2023).
- Rootmotion (2019). *FinalIK*. Available at: <http://root-motion.com/.Version2.1> (Accessed November, 2023).
- Rossen, B., Lind, S., and Lok, B. (2009). "Human-centered distributed conversational modeling: efficient modeling of robust virtual human conversations," in *Proceedings of the 9th international conference on intelligent virtual agents* (Springer-Verlag), 474–481. doi:10.1007/978-3-642-04380-2_52
- Roth, D., and Latoschik, M. E. (2020). Construction of the virtual embodiment questionnaire (veq). *IEEE Trans. Vis. Comput. Graph.* 26, 3546–3556. doi:10.1109/TVCG.2020.3023603
- Rzayev, R., Karaman, G., Wolf, K., Henze, N., and Schwind, V. (2019). "The effect of presence and appearance of guides in virtual reality exhibitions," in *Proceedings of Mensch und Computer 2019* (New York, NY, USA: Association for Computing Machinery). MuC '19. doi:10.1145/3340764.3340802
- Salagean, A., Crellin, E., Parsons, M., Cosker, D., and Stanton Fraser, D. (2023). "Meeting your virtual twin: effects of photorealism and personalization on embodiment, self-identification and perception of self-avatars in virtual reality," in *Proceedings of the 2023 CHI conference on human factors in computing systems*, 1–16. doi:10.1145/3544548.3581182
- Scavarelli, A., Arya, A., and Teather, R. J. (2021). Virtual reality and augmented reality in social learning spaces: a literature review. *Virtual Real.* 25, 257–277. doi:10.1007/s10055-020-00444-8
- Schroeder, R. (2002). *Social interaction in virtual environments: key issues, common themes, and a framework for research*. London: Springer, 1–18. doi:10.1007/978-1-4471-0277-9_1
- Schubert, T., Friedmann, F., and Regenbrecht, H. (2001). The experience of presence: factor analytic insights. *Presence Teleoperators Virtual Environ.* 10, 266–281. doi:10.1162/105474601300343603
- Schulz, R. (2023). Welcome to the metaverse: a comprehensive list of social VR/AR platforms and virtual worlds. Available at: <https://ryanschultz.com/list-of-social-vr-virtual-worlds/> (Accessed November, 2023).
- Skarbez, R., Brooks, F. P., Jr., and Whitton, M. C. (2017a). A survey of presence and related concepts. *ACM Comput. Surv.* 50, 96:1–96:39. doi:10.1145/3134301
- Skarbez, R., Brooks, F. P., and Whitton, M. C. (2021). Immersion and coherence: research agenda and early results. *IEEE Trans. Vis. Comput. Graph.* 27, 3839–3850. doi:10.1109/TVCG.2020.2983701
- Skarbez, R., Neyret, S., Brooks, F. P., Slater, M., and Whitton, M. C. (2017b). A psychophysical experiment regarding components of the plausibility illusion. *IEEE Trans. Vis. Comput. Graph.* 23, 1369–1378. doi:10.1109/TVCG.2017.2657158
- Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Trans. R. Soc. B Biol. Sci.* 364, 3549–3557. doi:10.1098/rstb.2009.0138
- Slater, M., Banakou, D., Beacco, A., Gallego, J., Macia-Varela, F., and Oliva, R. (2022). A separate reality: an update on place illusion and plausibility in virtual reality. *Front. Virtual Real.* 3. doi:10.3389/frvir.2022.914392
- Slater, M., Pérez Marcos, D., Ehrsson, H., and Sanchez-Vives, M. V. (2009). Inducing illusory ownership of a virtual body. *Front. Neurosci.* 3, 214–220. doi:10.3389/neuro.01.029.2009
- Slater, M., Spanlang, B., Sanchez-Vives, M. V., and Blanke, O. (2010). First person experience of body transfer in virtual reality. *PLOS ONE* 5, e10564–e10569. doi:10.1371/journal.pone.0010564
- Slater, M., and Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. *Presence Teleoperators Virtual Environ.* 6, 603–616. doi:10.1162/pres.1997.6.6.603
- Stauffert, J.-P., Niebling, F., and Latoschik, M. E. (2020). Latency and cybersickness: impact, causes, and measures. A review. *Front. Virtual Real.* 1. doi:10.3389/frvir.2020.582204
- Stoll, C., Hasler, N., Gall, J., Seidel, H.-P., and Theobalt, C. (2011). "Fast articulated motion tracking using a sums of Gaussians body model," in *2011 international conference on computer vision*, 951–958. doi:10.1109/ICCV.2011.6126338
- The Capture (2021). *Capture live*. Available at: <https://capture.com.Version254b> (Accessed November, 2023).
- Unity Technologies (2020). *Unity*. Available at: <https://unity.com.Version2020.3.25f1> (Accessed November, 2023).
- Valve Corporation (2021). *SteamVR*. Available at: <https://store.steampowered.com/app/250820/SteamVR> (Accessed November, 2023).
- Valve Corporation (2022). *Index*. Available at: <https://store.steampowered.com/vliveindex> (Accessed November, 2023).
- Volante, M., Babu, S. V., Chaturvedi, H., Newsome, N., Ebrahimi, E., Roy, T., et al. (2016). Effects of virtual human appearance fidelity on emotion contagion in affective inter-personal simulations. *IEEE Trans. Vis. Comput. Graph.* 22, 1326–1335. doi:10.1109/TVCG.2016.2518158
- Volante, M., Hsu, Y.-C., Liu, K.-Y., Mazer, J. P., Wong, S.-K., and Babu, S. V. (2020). "Effects of interacting with a crowd of emotional virtual humans on users' affective and non-verbal behaviors," in *2020 IEEE conference on virtual reality and 3D user interfaces (VR)*, 293–302.
- Volante, M., Robb, A., Duchowski, A. T., and Babu, S. V. (2018). "Empirical evaluation of virtual human conversational and affective animations on visual attention in inter-personal simulations," in *2018 IEEE conference on virtual reality and 3D user interfaces (VR)*, 25–32. doi:10.1109/VR.2018.8446364
- Waltemate, T., Gall, D., Roth, D., Botsch, M., and Latoschik, M. E. (2018). The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE Trans. Vis. Comput. Graph.* 24, 1643–1652. doi:10.1109/TVCG.2018.2794629
- Weidner, F., Boettcher, G., Arboleda, S. A., Diao, C., Sinani, L., Kunert, C., et al. (2023). A systematic review on the visualization of avatars and agents in AR and VR displayed using head-mounted displays. *IEEE Trans. Vis. Comput. Graph.* 29, 2596–2606. doi:10.1109/TVCG.2023.3247072
- Wenninger, S., Achenbach, J., Bartl, A., Latoschik, M. E., and Botsch, M. (2020). "Realistic virtual humans from smartphone videos," in *Proceedings of the 26th ACM symposium on virtual reality software and technology* (New York, NY: Association for Computing Machinery). VRST '20. doi:10.1145/3385956.3418940
- Wienrich, C., Döllinger, N., and Hein, R. (2021). Behavioral framework of immersive technologies (BehaveFIT): how and why virtual reality can support behavioral change processes. *Front. Virtual Real.* 2. doi:10.3389/frvir.2021.627194
- Wienrich, C., and Latoschik, M. E. (2021). Extended artificial intelligence: new prospects of human-ai interaction research. *Front. Virtual Real.* 2. doi:10.3389/frvir.2021.686783

- Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., et al. (2007). A process model of the formation of spatial presence experiences. *Media Psychol.* 9, 493–525. doi:10.1080/15213260701283079
- Wisessing, P., Dingliana, J., and McDonnell, R. (2016). “Perception of lighting and shading for animated virtual characters,” in *Proceedings of the ACM symposium on applied perception* (New York, NY, USA: Association for Computing Machinery), 25–29. doi:10.1145/2931002.2931015
- Wolf, E., Döllinger, N., Mal, D., Wenninger, S., Andrea, B., Botsch, M., et al. (2022a). Does distance matter? Embodiment and perception of personalized avatars in relation to the self-observation distance in virtual reality. *Front. Virtual Real.* 3. doi:10.3389/frvir.2022.1031093
- Wolf, E., Döllinger, N., Mal, D., Wienrich, C., Botsch, M., and Latoschik, M. E. (2020). “Body weight perception of females using photorealistic avatars in virtual and augmented reality,” in *2020 IEEE international symposium on mixed and augmented reality (ISMAR)*, 462–473. doi:10.1109/ISMAR50242.2020.00071
- Wolf, E., Fiedler, M. L., Döllinger, N., Wienrich, C., and Latoschik, M. E. (2022b). “Exploring presence, avatar embodiment, and body perception with a holographic augmented reality mirror,” in *2022 IEEE conference on virtual reality and 3D user interfaces (VR)*, 350–359. doi:10.1109/VR51125.2022.00054
- Wolf, E., Mal, D., Frohnapfel, V., Döllinger, N., Wenninger, S., Botsch, M., et al. (2022c). “Plausibility and perception of personalized virtual humans between virtual and augmented reality,” in *2022 IEEE international symposium on mixed and augmented reality (ISMAR)*, 489–498. doi:10.1109/ISMAR55827.2022.00065
- Zell, E., Aliaga, C., Jarabo, A., Zibrek, K., Gutierrez, D., McDonnell, R., et al. (2015). *To stylize or not to stylize? the effect of shape and material stylization on the perception of computer-generated faces* 34. doi:10.1145/2816795.2818126
- Zell, E., Zibrek, K., and McDonnell, R. (2019). “Perception of virtual characters,” in *ACM SIGGRAPH 2019 courses* (New York, NY, USA: Association for Computing Machinery). SIGGRAPH '19. doi:10.1145/3305366.3328101
- Zibrek, K., Kokkinara, E., and McDonnell, R. (2017). “Don’t stand so close to me: investigating the effect of control on the appeal of virtual humans using immersion and a proximity-based behavioral task,” in *Proceedings of the ACM symposium on applied perception*, 1–11. SAP '17. doi:10.1145/3119881.3119887
- Zibrek, K., Martin, S., and McDonnell, R. (2019). Is photorealism important for perception of expressive virtual humans in virtual reality? *ACM Trans. Appl. Percept.* 16, 14:1–19. doi:10.1145/3349609