A Systematic Review on the Visualization of Avatars and Agents in AR & VR displayed using Head-Mounted Displays

Article *in* IEEE Transactions on Visualization and Computer Graphics · May 2023

DOI: 10.1109/TVCG.2023.3247072

Florian Weidner
Lancaster University

Stephanie Arévalo Arboleda
Technische Universität Ilmenau

Chenyao Diao

Christian Kunert
Technische Universität Ilmenau

A Systematic Review on the Visualization of Avatars and Agents in AR & VR displayed using Head-Mounted Displays

Florian Weidner, Gerd Boettcher, Stephanie Arevalo Arboleda, Chenyao Diao, Luljeta Sinani, Christian Kunert, Christoph Gerhardt, Wolfgang Broll, Alexander Raake

Abstract— Augmented Reality (AR) and Virtual Reality (VR) are pushing from the labs towards consumers, especially with social applications. These applications require visual representations of humans and intelligent entities. However, displaying and animating photo-realistic models comes with a high technical cost while low-fidelity representations may evoke eeriness and overall could degrade an experience. Thus, it is important to carefully select what kind of avatar to display. This article investigates the effects of rendering style and visible body parts in AR and VR by adopting a systematic literature review. We analyzed 72 papers that compare various avatar representations. Our analysis includes an outline of the research published between 2015 and 2022 on the topic of avatars and agents in AR and VR displayed using head-mounted displays, covering aspects like visible body parts (e.g., hands only, hands and head, full-body) and rendering style (e.g., abstract, cartoon, realistic); an overview of collected objective and subjective measures (e.g., task performance, presence, user experience, body ownership); and a classification of tasks where avatars and agents were used into task domains (physical activity, hand interaction, communication, game-like scenarios, and education/training). We discuss and synthesize our results within the context of today's AR and VR ecosystem, provide guidelines for practitioners, and finally identify and present promising research opportunities to encourage future research of avatars and agents in AR/VR environments.

Index Terms—Virtual reality, augmented reality, avatars, visualization.

1 Introduction

Augmented Reality (AR) and Virtual Reality (VR) have been around for decades. However, the availability of affordable consumer hardware has allowed a wider audience to experience immersive environments. Here, many popular AR/VR applications use avatars or agents. While an avatar is controlled by a human, an agent is controlled solely by a computer system [73]. Both can be defined as "perceivable digital representations" [73] that can interact with other virtual or real elements in the environment. However, the nature of the interaction can determine the importance of certain aspects such as rendering style [49] or showing only certain body parts [27].

Previous works have already reviewed different aspects of avatars and agents. McDonnell [108] presented a seminal overview of various aspects such as gender representation, conversational motion, locomotion, body type, and rendering style outside of AR and VR. In 2016, Hantono et al. [90] investigated agents for AR in the domain of education. Similarly, Hudson & Hurter [95] reviewed avatar types and their applications in sports, education, and military. In addition, Green et al. [88] presented a review of 18 papers that focus on the use of avatars in gaming. Norouzi et al. [110] presented a systematic review of interactive, embodied AR agents, viewed using head-mounted displays (HMDs), to improve the understanding of interactive agents in AR environments as well as their utilization in specific applications. Other authors have presented reviews on topics closely related to avatars and agents: Katsyri et al. [99] review the state of the art with respect to the uncanny valley. Oh et al. [111] focus on the investigation of social presence. Genay et al. [84] provide a review on embodiment and Korban & Li [104] present a survey on avatars in shared virtual environments.

In this work, we go beyond reviewing domain-specific studies. Our goal is to collect and compare the different visualization styles and body types that have been used for agents and avatars, displayed through HMDs. Additionally, we outline how they influenced perception, task performance, sense of agency, social presence, or embodiment in different studies. Thus, our core contributions are:

1. Classification of studies using avatars and agents into task domains.
2. An analysis of the effects of rendering styles and body parts.
3. General guidelines for the implementation of avatars and agents.
4. Future research directions.

The paper is structured as follows. Sect. 2 describes the methodology of our systematic literature review including the search process. Sect. 3 presents an overview of the research papers used in our analysis, including study design, hardware, avatars and agents, and investigated tasks. In Sect. 4, we explore the results of the analyzed papers sorted by task domain. In Sect. 5, we discuss our findings and propose guidelines for avatar and agent design and deployment. Finally, Sect. 6 summarizes our paper and provides a brief outlook.

2 Materials and methods: PRISMA

For this article, a systematic literature review was conducted following the PRISMA method [117].

2.1 Step 1: Define review goal

In this article, we provide an overview of studies that investigate the effects of different avatar or agent visualizations in AR and VR , focusing on HMD-based setups. We target articles that compare at least two types of representations. By that, we aim to provide a systematic overview on the state of the art and identify and categorize different types of avatars and agents along with the tasks where they were used.

2.2 Step 2: Identifying search terms and databases

Our search terms can be summarized as follows:

(*"virtual reality"* OR *"augmented reality"* OR *"virtual environment"*) AND (*"visualization"* OR *"fidelity"* OR *"appearance"* OR *"render*"*) AND (*"avatar"* OR *"agent"* OR *"virtual human"*)

The search terms consider topic areas of AR and VR, avatars and agents, as well as domain-specific terms. We added domain-specific terms (visualization, fidelity, appearance, and rendering ("render*"))

- *Florian Weidner (florian.weidner@tu-ilmenau.de, Gerd Boettcher (gerd.boettcher@tu-ilmenau.de), Christian Kunert (christian.kunert@tu-ilmenau.de), Christoph Gerhardt (christoph.gerhardt@tu-ilmenau.de), and Wolfgang Broll (wolfgang.broll@tu-ilmenau.de) are with the Virtual Worlds and Digital Games Group, Technische Universität Ilmenau, Germany.*
- *Stephanie Arevalo Arboleda (stephanie.arevalo@tu-ilmenau.de), Chenyao Diao (chenyao.diao@tu-ilmenau.de), Luljeta Sinani (luljeta.sinani@tu-ilmenau.de), and Alexander Raake (alexander.raake@tu-ilmenau.de) are with the Audiovisual Technology Group, Technische Universität Ilmenau, Germany.*

to focus on articles that focus on rendering styles, ignoring tracking or avatar generation methods.

We searched in the digital libraries of the Association for Computing Machinery (ACM Digital Library), the Institute of Electrical and Electronics Engineers (IEEExplore), Springer Link, Frontiers, and MIT Press. We selected these databases because they cover major conferences like IEEE ISMAR, IEEE VR, ACM SIGCHI, but also journal articles like IEEE TVCG.

2.3 Step 3: Search process

We searched title, abstract, and keywords, and included all work published between 2015 and 2022. This time frame was chosen to focus on studies that were conducted after the widespread availability of HMDs (i.e., the introduction of the Oculus Rift Developer Kit in 2014 and the release of the HTC Vive in 2016). While earlier research in this domain exists and has been covered by earlier reviews (cf. Sect. 1), these new generation of devices provided a rather good technical visual quality that fostered research on avatars and agents. We started the first iteration of the search process on April 26th and finished it on May 12th, 2021. After removing duplicates, the search process delivered a list of 2772 papers. We performed the second iteration covering articles from May 2021 to August 2022 on September 6th, 2022, It resulted in 447 additional papers. Furthermore, we added articles based on our expertise and prior reviews.

2.4 Step 4: Judging quality and relevancy

We limited the search to the respected and well-known libraries of IEEE, ACM, Springer, Frontiers, and MIT Press. Therefore, we did not discard articles due to a lack of quality.

In each iteration, we analyzed the titles and abstracts to determine if the articles focus on avatars or agents, and AR or VR. Additionally, we evaluated the papers by answering to the following questions:

- Is the article about avatars or agents in AR or VR?
- Does it use a head-mounted display?
- Does the article include a user study and evaluation?
- Does the evaluation feature at least two conditions?

Only papers that did use an HMD during the study were considered relevant. We did this as HMDs are (right now) the popular devices for immersive applications. Papers that did not mention AR or VR were excluded. We also discarded papers that did not include a user study to ensure that the analyzed studies featured reproducible experiments and comparisons based on criteria relevant to human users. Also, if the paper did not feature a comparison between at least two conditions, we removed it from the final selection. Accurate numbers on the filter process are not available as we excluded the papers based on the first reason we found. After that, we did not check the other criteria.

This exclusion process resulted in 47 studies that matched all the above mentioned criteria. These 47 papers were further analyzed in an in-depth review process. We also applied backtracking on those 47 papers. Here, we checked all publications referenced by those papers and used the same evaluation criteria. This added 25 articles, resulting in a total of 72 articles.

2.5 Limitations

Although our review process has been handled as precise and systematic as possible, we identified some limitations and concerns connected to the methods. The first one is related to the selection of the databases. IEEE, ACM and SpringerLink cover a large number of publications, but it is still possible that important research on this topic was not found simply because it was listed in different databases (e.g., MDPI or Elsevier). Also, it is possible that some relevant papers were missed as they did not use one of our search terms in their keywords, title, or abstract. For instance, the term "mixed reality" could have been used instead of "virtual reality" and "augmented reality". We could have also added further search terms like "evaluation", "preference", "experience" or "perception" to expand the search. We tried to mitigate this issue via backtracking. We also excluded CAVEs and other projection-based systems to focus on the prominent domain of HMD-based AR and VR. For future researchers, it is important to also investigate if our findings transfer to non-HMD-based systems as well as to non-AR and non-VR applications such as video games.

Our research objective, search terms, and exclusion criteria focus on the question of how to display avatars or agents, with respect to different rendering styles and different body parts. Therefore, we did not include articles that investigate only behavioral/tracking fidelity or realism (e.g., Gonzalez-Franco et al. [87], Bailenson et al. [74] or Oh Kruzic et al. [112]). Similarly, we did not investigate differences within a single rendering styles (e.g., Hasler et al. [92], Banakouk et al. [76] or Piryankova et al. [118]). We acknowledge that these are important aspects. Both cover a variety of facets, e.g., facial animations, body movements, quality of tracking devices, quality of animations, skin-color matching, influence of social biases (race, intelligence, age), and many more. However, we decided to exclude those aspects and to focus on our core question about rendering style and body parts.

3 Overview of collected articles

In this section, we provide an overview of the 72 chosen articles. We report on the study design and type, e.g., collected data and measures, tasks used, sample description; technical information, e.g., displays, software, and — most important — the characteristics of the used agents and avatars. The oldest paper in this review was published in 2015 and the newest one in 2022. Four articles are from 2015, five from 2016. The time period from 2017 to 2021 has approximately 12 papers per year and we have six papers from 2022. Most of the papers (65) focus on research on VR. While 6 focus on AR only 1 article deals with an asymmetric AR/VR setup.

3.1 Study design

In our corpus, 45 papers used a within-subjects design, 12 papers used a between-subjects design, and 15 papers used a mixed-factorial design. The mean age of the participants in all studies was 24.4 years (SD = 4.1, median = 23.7, min = 7.4, max = 36.5) (note: 14 articles did not provide information about age). The number of participants ranges from 7 – 1106 with a median of 25 (Avg. = 66, SD = 156.6). On average, 35.8% of the participants self-identified as male, 25.6% as female, and 0.1% as other. Regarding the type of measures considered, 40 papers collected both quantitative and qualitative data, 30 collected only qualitative data, and two collected only quantitative data.

3.2 Measures

The measures that were used in the majority of the papers were: task performance, measured in 24 (33.3%) papers; presence in 22 (30.6%) papers; body ownership in 19 (26.4%) papers, and social presence in 13 (18.1%) papers. These were followed by communication behavior in 12 (16.7%) papers and agency in 11 (15.3%) papers. Other measures include simulator sickness in 10 (13.9%) papers, co-presence in nine (12.5%), preference in eight (11.1%), embodiment in seven (9.7%), physiological parameter in six (8.3%), system usability and user experience in five (6.9%), uncanny valley, enjoyment and cognitive load in four (5.6%) and spatial presence in two (2.8%).

Among the formalized questionnaires used in the studies are several related to presence (Slater-Usoh-Steed Presence Questionnaire [125], Igroup Presence Questionnaire (IPQ) [96], Witmer and Singer Presence Questionnaire [129], Networked Minds Social Presence Inventory [77]). Embodiment and body ownership were often measured with the Illusion of Virtual Body Ownership Questionnaire [121] and the Embodiment Questionnaire based on Gonzalez–Franco & Peck [115].

Next to that, the authors often used the Uncanny Valley Questionnaire [94], the Immersive Experience Questionnaire [120], and the Interpersonal Communication Questionnaire [122]. Other prominent instruments were the AttrakDiff Questionnaire [93], the System Usability Scale [79], the Simulation Sickness Questionnaire [100], and the NASA Task Load Index (TLX) [91].

3.3 Avatars and agents

3.3.1 Rendering styles

The rendering style classifications were derived in the following way: Each author surveyed a batch of the 72 articles and described the rendering style in a few words. Later, in a joint effort, the authors refined those descriptions and inductively created and assigned the labels. *Abstract* visualizations (n = 21, 29.2%) remotely resemble human appearance. Instead, they consist only of a single or very few basic geometric shapes like a cylinder or a cuboid. Fig. 1a illustrates such avatars. Another simple design is the *stickman* (n = 10, 13.9%). This design resembles the human bone structure used to animate 3D characters. It does not include details such as skin or clothing and often features few colors (see Fig. 1f). Three papers (4.2%) feature a *hybrid* design that includes a real video. One example is a head which is seen on a screen that is floating on top of the body of a robot avatar (see Fig. 1g). A different design is the *robot* (n = 17, 23.6%) which also represents human-like proportions in combination with a mechanical look and often metallic textures (see Fig. 1d). A *cartoon* design was used in 16 studies (22.2%). This design does not have realistic human proportions, instead it implements exaggerated body parts like a bigger head, nose, ears, and hands. It features non-typical elements, e.g., four fingers per hand (see Fig. 1e). The *stylized* (n = 20, 27.8%) design features human proportions with detailed body parts as well as visible facial characteristics like a nose and ears. It looks very similar to a realistic 3D model of a human but without any human-like textures and does not necessarily follow human morphology (see Fig. 1h). Major difference between *cartoon* and *stylized* is that the former does not follow human morphology (e.g., it can have only four fingers).

Five different studies (6.9%) implemented a *video* avatar. This design requires the user, for example, to be filmed in front of a green screen and the extracted video feed appears as a two-dimensional avatar in the virtual surrounding (see Fig. 1i). *Point cloud* was chosen by six studies (8.3%). This version represents the user as a cloud of single dots. The overall shape of the cloud mimics the proportions of a human, although the level of detail depends on the number and color of the dots in the cloud (see Fig. 1c). While the right image is not a traditional point cloud, we classified it here as the authors named it as such. It could also belong to the category *abstract*. A *translucent* avatar was included in three studies (4.2%). Those designs are presented as detailed 3D models but are not fully opaque (see Fig. 1j).

Several articles implemented different kinds of realistic human visualizations: generic realistic, personalized, and scanned. The *generic realistic* representation is featured in 56 papers (77.8%). This style is a very detailed model of a human with textures and clothing created with character modeling software like Autodesk Maya (https://www.autodesk.de/products/maya/) or Blender (https://www.blender.org/). However, it does not resemble the participants in the study. The *scanned realistic* style (n = 3, 4.2%) on the other hand is based on a real person using photogrammetry and/or depth cameras. The *personalized realistic* style (n = 10, 13.9%) has the same level of fidelity as the generic realistic one but is also based on the actual participant. It is created, for example, by capturing images of the participants with cameras and then combining these scans with preexisting 3D models (see Fig. 1b).

3.3.2 Body parts

Independent from the visual fidelity, the visual representation differs in how much of the avatar's or agent's body is visible. Overall, the studies in our corpus evaluated 15 different types of body part configurations. In 13 papers (18.1%), no visualization of the body was chosen at all in at least one condition (often as a control). These conditions did not show any visualization (neither hands nor other parts of the skeleton). Two papers (2.8%) used an abstract body representation (a single cube [57] and a spherical drone [39]). There are several studies featuring only some body parts: hands-only (n = 20, 27.8%), hands+arms (n = 9, 12.5%), hands+feet (n = 3, 4.2%), hands+head (n = 2, 2.8%), head+torso (n = 2, 2.8%), hands+torso (n = 2, 2.8%), head+hands+torso (n = 2, 2.8%), hands+head+torso (n = 2, 2.8%), hands+arms+legs (n = 1, 1.4%), hands+head+feet (n = 1, 1.4%), hands+forearms (n = 1, 1.4%), hands+feet+head+arms (n = 1, 1.4%), head-to-knee-body (n = 1, 1.4%), feet-only (n = 1, 1.4%). 50 papers (69.4%) used full-body visualizations in at least one condition. Similarly to the visualizations, researchers investigated a wide variety of body part configurations, motivating a structured synthesis of the existing literature to provide guidelines.

3.4 Task domains

We classified the studies into the following task domains, using an iterative inductive approach during the analysis, similar to the procedure described in Sect. 3.3.1:

1. Physical activity: Walking
Non-walking
2. Hand interaction: Direct hand interaction
Indirect hand interaction
3. Communication: Dynamic/bi-directional
Static/uni-directional
4. Game-like scenario
5. Education/training

Physical activity includes tasks that revolve around moving and using the entire body, e.g., walking, fitness, or bow shooting. These tasks required participants to stand during the experiment. As many of the activities focused on walking, we split the articles into *walking* and *non-walking*.

Hand interaction contains studies that focus on actions that are performed solely with the hands, like pointing, touching, or selecting buttons on a tablet. Because the hands were the only body parts that many studies exclusively focused on, we decided to put them in a category of their own. In *direct hand interaction*, the tasks mainly revolve around 3D manipulation of virtual objects. Opposed to this, the tasks in *indirect hand interaction* focus more on feeling, simple touching, pointing, or interacting with real objects using passive haptics.

Communication groups experiments that deal with communication scenarios. *Dynamic/bi-directional communication* features studies where users freely talk to an actor, another participant, the experimenter, or an intelligent agent. In *static/uni-directional communication*, conversations are either unidirectional (plain listening) or one users can only use a limited set of questions with an agent. We divided these categories of communication because we consider the task of having a "natural" conversation to have different or higher requirements on rendering style and visible body parts compared to a conversation where a participant only listens to an agent. The category *game-like scenario* incorporates papers where the task and setting is similar to modern video games. These studies involve moving through and interacting with the environment as well as acting in a stressful or time-sensitive scenario or bow-shooting. The category *education/training* contains articles where the activity of the experiment was centered on knowledge and skill transfer. While these papers could also be in communication or physical activity, the specific nature of learning and knowledge retention encouraged us to group them in a dedicated category. Some studies would have qualified for more than one task domain, or the distinction is at least debatable. We will mention such examples in the following sections.

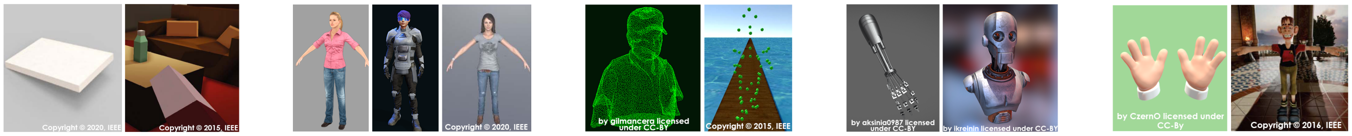
4 Analysis by Task Domains

4.1 Physical activity

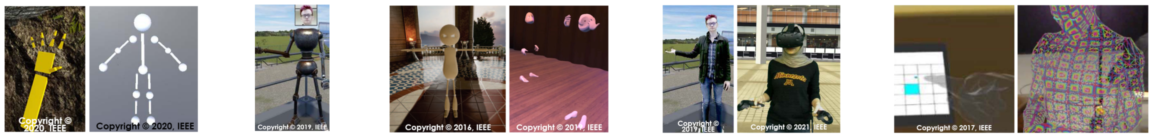
4.1.1 Walking

Seven VR articles belong to the category physical activity - walking. Here, five papers evaluated self-avatars [18, 40, 44, 59, 63] and two evaluated agents [9, 69].

The tasks required participants to walk around a scene, observe walking avatars, or to walk and avoid obstacles. In the experiment conducted by Valkov et al. [59], participants were supposed to walk towards markers in VR. Matsuda et al. [40] as well as Choi et al. [9] let participants mainly observe avatars in first- as well as third-person and



(a) Abstract - left: [68], right: [37]
(b) Realistic - left: generic cf. [11], middle: personalized cf. [22], scanned: right [12]
(c) Point cloud - left: cf. [13], right: [63]
(d) Robot - cf. [53], [49]
(e) Cartoon - left: cf. [53], right: [59]



(f) Stickman - left: [35], right: [12]
(g) Hybrid [64]
(h) Stylized - left: [59], right: [16]
(i) Video - left: [64], right: [4]
(j) Translucent - left: [1], right: [47]

Fig. 1: Avatar and agent types (snippets from the original research papers and publicly available models for illustrative and explanatory purposes).

answer questionnaires. Zibrek et al. [69] realized avatar observation by placing one label on the front and one on the back of the avatar and instructing participants to read them. In a study done by Wirth et al. [63], participants walked on a treadmill while wearing a VR HMD. Ogawa et al. [44] created virtual rooms where subjects needed to walk to buttons and press them. Finally, Gorisse et al. [18] set up various tasks where subjects are required to traverse increasingly difficult environments and dodge obstacles.

All studies included generic realistic avatars in their experiments with the exception of Gorisse et al. [18] who relied on scanned realistic avatars. Only Wirth et al. [63] included a point-cloud-based avatar. Valkov et al. [59] and Zibrek et al. [69] added avatars that are visualized in a cartoon style. Ogawa et al. [44] and Gorisse et al. [18] included robot styles. Choi et al. [9] as well as Matsuda et al. [40] did only include a generic realistic style and focused on avatar body parts. With regards to the avatar type, all seven studies include full-body avatars in their experiments. Ogawa et al. [42] and Matsuda et al. [40] also included hands-only or hands+feet, respectively. An avatar from head to knees was included by Choi et al. [9].

The analyzed papers generally point to the conclusion that full-body avatars and agents as well as higher visual fidelity can positively influence several aspects like presence [44], user experience [63], realism [69], task performance [59], body ownership [18, 40], and agency [69]. A scanned self-avatar may also increase the sense of embodiment but only one of the studies included such an avatar style [18]. With a heightened sense of presence and embodiment that can be facilitated by full-body or self-representing avatars, subjects are less willing to show risk-taking behavior [18] or walk through walls [44]. Valkov et al. [59] found that a realistic visualization was not important for estimating distances. Ogawa et al. [9, 44] found that, while a realistic visualization can lead to more realistic behavior of users (e.g., walking speed and stride), the sense of body ownership was actually rated higher for abstract avatars [44]. The authors attribute this to the uncanny valley effect in the case of full-body avatars. For the hands-only avatar, they point out that the realistic hands model did not properly align with the real hands, leading to a higher sense of body ownership in the abstract-avatar condition. A scanned realistic self-avatar seems to lead to an increased sense of embodiment but special care should be taken during avatar creation so that affinity does not suffer [18].

In summary, literature for this task domain leans towards realistic and full-body self-avatars and agents. Subjects seem to be more averse to risk taking and generally show a more realistic behavior when they have more complete and realistic avatars. Scanned avatars may be an interesting aspect for future research as they are only represented by one study in this category but they also need further consideration regarding reconstruction problems and the uncanny valley effect. Also, none of the studies in this category consider AR systems which would therefore be an interesting topic to explore in this context.

4.1.2 Non-walking

This category consists of eight VR articles. Six evaluated self-avatars [5, 12, 22, 28, 29, 60], one an avatar of another person [66] and one an agent [7]. The tasks in this category require participants to move but the focus is not on walking. Fribourg et al. [12] used various fitness exercises (yoga, boxing, soccer) and Waltemate et al. [60] simple body movements. One study required participants to follow a historical repair scenario [7]. Other experiments used simpler tasks, like high-fiving another character [66], freely interacting in a room and in front of a mirror [29], stepping of a virtual ledge [5], and reacting to certain events [28]. Following movements was part of two studies [22, 60]. Six studies observed the influence of different visualization styles [5, 7, 12, 22, 29, 60]. Here, all studies except two [28, 29] used one or more realistic self-avatars (personalized, generic, and/or scanned). Jo et al. [22] used a personalized realistic but stylized avatar (similar to https://readyplayer.me/) — which is unique in our corpus. Other rendering styles are abstract, robot, cartoon, none, and stickman which covers most of our categories. All experiments used a full-body avatar. Two articles additionally investigated the effect of visible body parts [28, 66]. These studies had hands-only [66] and hands+feet [28] as additional conditions.

The results show that body ownership benefits from a personalized style. These personalized self-avatars were either realistic [60] or stylized [22]. Fribourg et al. [12] report that embodiment is better with a personalized realistic self-avatar. However, this contradicts the results of Krogmeier & Mousas [29] where embodiment was not affected by the rendering style. An similar contradiction happens with presence. Jo et al. [22] report no difference in their comparison of personalized realistic, scanned realistic, and generic realistic, whereas a personalized realistic avatar led to more presence than a generic realistic self-avatar in the experiment of Waltemate et al. [60]. A reason for this could be the rather unrealistic setting (a cave-like temple) as well as missing computer graphic effects like shadows in the experiment by Jo et al. [22], which might have led to breaks in presence and lower overall presence. Young et al. [66] show that full-body avatars lead to higher presence than a hands-only version. For agents, Butz et al. [7] report no difference in presence when comparing a realistic agent to an abstract one. Other noteworthy findings are the increased emotional response when using personalized realistic avatars compared to generic realistic ones [60]. Further, Krogmeier & Mousas [29] report a relationship between eye fixations and embodiment as well as a gender effect regarding agency. Kondo et al. [28] found no differences in proprioception — regardless of avatar or agent configuration or style.

The difference in presence — despite similar tasks — that we observed between Jo et al. [22] and Waltemate et al. [60] motivates further research. One potential reason for the difference is that participants of the former study saw an avatar whose face was hidden behind the 3D model of an HMD. This was not the case in the latter study. For embodiment, the nature of the task might be important. In Krogmeier & Mousas' experiment [29], the task was fully focused on the avatar and the rendering style did not lead to differences in embodiment. The task in the experiment performed by Fribourg et al. [12] required the participants to focus not only on the body but also on the correct execution of movements. Here, the authors report differences in embodiment.

Overall for this task domain, a generic realistic or even a personalized realistic full-body self-avatar could be promising. Reported results with this configuration show improvements in attractiveness, body ownership, embodiment, emotional response, and presence. Further investigations of various realistic configurations may pinpoint causes for the inconsistent results.

4.2 Hand interaction

4.2.1 Direct hand interaction

This task domain includes 10 papers with tasks that are related to hand interaction of self-avatars. The activities in this domain include reaching [43, 58] and selecting as well as manipulating objects [3, 10, 23, 27, 33, 51, 52, 68]. The main goal of these papers was to evaluate different hand representations when interacting with elements in the environment. The avatars' appearance included showing only the hand [23, 33, 43] and hand+forearm [3, 10, 27, 51, 52, 68]. The styles were stylized [23], generic realistic [3,10,27,33,43,51,52,58,68], transparent [10], abstract [3, 23, 33, 33, 43, 51, 52], iconic [3, 27, 51, 52, 68], and robot [51, 52, 68].

Regarding the sense of body ownership, seven out of nine papers found that realistic avatar hands [3, 10, 33, 43, 68] or stylized hands [23] generate a higher sense of ownership or embodiment. Closely related to embodiment, Kocur et al. [27] compared the perception of phantom pain when using realistic hand representations and abstract ones. Here, the results point to higher phantom pain for realistic hands. In terms of agency, there were mostly no significant differences. Only Argelaguet et al. [3] found a stronger sense of agency for abstract hand representations, while in Dewez et al. [10] a stronger feeling of control was found for realistic representations in direct manipulation. Similarly, there were mostly no significant differences in task performance. However, when significant differences were found they point to specific task conditions, e.g., the preference of abstract hands in precise pointing [33] or the effect of missing fingers in frequency of finger use [27]. Additionally, four studies [27, 51, 52, 68] that evaluated presence showed contradicting results for hand representations with respect to task performance. Interestingly, Schwind et al. [52] found differences between embodied male and female hands. Male participants reported a higher level of presence when using realistic hands, while female participants reported more presence with abstract or cartoon-like hands. Also, female participants reported lower levels of presence with male hands, whereas such an effect was not observed for the male participants.

In this task domain, the hand representations were mostly realistic hands in comparison to abstract versions. The results show a higher sense of body ownership when using realistic hands. Also, most studies showed only the hand and the forearm during the task. We attribute this to the nature of the tasks used, i.e., manipulation tasks. Here, some studies compared showing different parts of the arm [58] or the effect of missing fingers [27]. Further, we highlight the different results shown in terms of task performance, agency, and presence. We attribute these differences to the difficulty of the task (e.g., the presence of obstacles [3, 58]) and the required resources (e.g., spatial planning and understanding may have a higher influence than the hand representation on preferences and performance).

4.2.2 Indirect hand interaction

Ten articles investigate indirect hand interactions in VR. Nine with self-avatars and one paper [19] evaluated both a self-avatar and the avatar of another person. They cover three sub-groups: interaction with real objects [1, 19, 26, 30, 53], reaction to unforeseen events [24, 25, 32], and pointing/perceiving objects [42, 54]. For the first group, interaction with real objects, three studies had participants type on real keyboards (passive haptics) [1, 19, 26], one study let them feel textures [53], and in another one, they had to touch their hand and foreign objects [30]. In the second group, unforeseen events, participants had to perform a distractor task and then react to a dangerous situation (a virtual knife cutting their hand [32] and a virtual stone dropping on their hand [24, 25]). In the third group, pointing/perceiving, participants had to point to objects [54] or estimate the size of cubes after interacting with them [42]. All but two articles focused on the rendering style of the hands. In these two, Jung et al. [24, 25] investigate generic realistic hands-only and hands+arms. They also modify participant's lower-body avatar (visible in a mirror; generic realistic and personalized realistic). Thus, these papers could also be in physical activity. Because the main task focuses on the hands, we decided to classify them in this category.

The investigated hand representations vary. In the interaction with real objects, the articles included translucent [1], generic realistic, [1,19,26,53], stickman [19], hybrid [30], video [19], robot [53], cartoon [53], abstract [26, 53], and point-cloud hands [53]. Knieriem et al. [26] also compared real hands and no hands. In the unforeseen-events-group, the comparisons include realistic (personalized [25], generic [24,25,32]), as well as robot [32], cartoon [32], and stickman hands [24]. The pointing/perceiving articles use generic realistic [42,54], robot [54], cartoon [54], stickman [42], and abstract hands [42, 54].

Results for the interaction-with-real-objects group are mixed. For the typing experiments, there was no best hand visualization. In different experiments, generic realistic and translucent ([26], none [1], abstract or none [19], or translucent hand [1, 26] performed best with respect to task performance. Experience seems to play a special role [26]. Similarly, other variables like preference [1, 19] and workload [19, 26] have various winners (translucent, video hand, generic realistic). These results suggest that, when typing on a real keyboard, various hand visualizations should be offered to users. For the article investigating texture perception [53], hand visualization did not influence task performance and among other measures (such as attractiveness, human-likeness, likability, eeriness), no clear favorite or best visualization emerged. In the experiments where participants had to react to events, visualizations that are a) generic realistic [32] and b) arms+hands lead to a higher virtual hand illusion than abstract, cartoon, or robot [32] and to higher presence and agency compared to only hands [25]. For pointing [53], abstract and robot hands led to the best task performance while generic realistic was most human-like. Here, the task of accurately pointing towards a target was better supported by the fine-grained limbs of the robot and abstract hands. Interestingly, robot hands were the most attractive. Cube size estimation was best with generic realistic hands (compared to abstract or stickman) [42].

Overall, the indirect hand interaction experiments in these studies show no clear results but favor generic realistic hands of self-avatars. Thus, a proper task analysis is recommended to truly find the most suitable hand visualization. In general, considering only task performance, either realistic (video, generic realistic) or task-appropriate hand visualizations (e.g., robot or abstract for pointing) perform best. Other variables like virtual hand ownership, body ownership, agency, workload, presence, or preference seem to most often benefit from realistic or video hand visualizations For simple interactions, abstract or robot hands seem to be appropriate and suitable. For more complex or longer interactions where presence, agency, or ownership is desired, more realistic hands seem to be more appropriate.

4.3 Communication

4.3.1 Dynamic/bi-directional communication

We classified 12 papers in this category [2, 4, 8, 13, 16, 20, 45, 56, 57, 61, 65, 67]. Two experiments [61, 65] were performed in AR, Yu et al. [67] conducted experiments as asymmetrical communication between AR and VR, whereas the rest of the articles conducted experiments in VR.

The studies in this category focus on scenarios where subjects actively interact and communicate with another subject that is visualized as an avatar. An exception is the study done by Wang et al. [61], where the communication is happening with a virtual agent. Similarly, in the majority of the studies, the subject of investigation was the conversation partner's visualization. However, Pan & Steed [45] and Heidicker et al. [20] explored the effect of self-avatar and Asseri et al. [4] evaluated both a self-avatar and the avatar of the conversation partner.

The tasks of these studies contain a variety of activities that participants had to solve together with another participant, with the experimenter, or with an intelligent agent. Two studies used a survival game [4, 20]. Five used puzzle-like activities [8, 13, 61, 65, 67]. Other tasks were conversation cards and playing charades [4], negotiation task and workspace manipulation [56], furniture movement [65], object manipulation [45], movement in front of a mirror [57], performing on stage [16] and having a discussion [2]. Some studies could also be in the category physical activity [45, 65] but since their main research objective was communication-focused, we included them in this category. Six out of 12 studies explored both rendering style and visible body parts [16, 20, 56, 57, 61, 65]. The other six studies [2, 4, 8, 13, 45, 67] focus on rendering style. Ten studies involve at least one condition with a full-body avatar. Cho et al. [8] and Agnew [2] used only upper-body avatars. These conditions are complemented by no visual representation at all [4, 56, 61], head+hands [20, 65], head+hand+feet [16], and head+torso [65]. Regarding rendering style, varieties of realistic styles were used. Scanned realistic style was used by Aseeri & Interrante [4], generic realistic style was used in 5 articles [13, 16, 45, 61, 65] and personalized realistic style was used in 3 articles [8, 57, 67]. Additional styles were realistic video avatars [4, 8], point-clouds [13, 67], volumetric capture [8], stylized [2], abstract [16, 57], and cartoon [2, 20, 65]. Two studies featured only one rendering style [56, 61] and four studies featured only one body part configuration [2, 8, 45, 67].

In all studies featuring varying body part visibility, the authors conclude that full-body visualizations outperform the other configurations. This include measures such as communication behavior [4, 13, 56, 61], social presence [4, 8, 20, 65], co-presence [8], presence [13], task performance [13, 45, 61], preference [4, 16, 61], and trust [4, 45]. Results of Yoon et al. [65] suggest that an upper-body can be a viable alternative to a full-body avatar depending on the task and collaboration context. For instance, communication with an avatar sitting behind a table requires only an upper-body. Looking at the influence of rendering styles on the evaluated measures, results vary widely. For instance, when social presence is evaluated, a real-video avatar [4], a point-cloud [13, 67], a cartoon style [20], a volume-capture [8], and a cartoon and a realistic avatar [65] performed best — depending on the experiment. Yoon et al. [65] mention that the rendering style (cartoon or realistic) had no influence on social presence but visibility of body parts had. Regarding communication behavior, results show that a generic realistic avatar [45, 61] and a video avatar [4] lead to a more human-like gaze behavior of the participant and more trust (even when compared to a scanned realistic avatar [4]). It is noteworthy that the level of avatar control is also relevant. A head+hands visualization performed equally well to full-body visualization in cases where no full-body tracking was available [20] (with respect to co-presence and behavioral interdependence). Gamelin et al. [13] report similar results and state that photo-realism is not necessarily as important as behavioral fidelity.

Overall, the studies suggest that a full-body avatar of another person is highly beneficial for bi-directional communication. Although realism seems to add value for communication behavior and task-performance, results are less conclusive with respect to social presence. Here, cartoon avatars performed also rather well.

4.3.2 Static/uni-directional communication

We found 12 articles that deal with a static communication scenario. Three of these perform their study in AR [21, 47, 62]. The rest operates in VR [6, 21, 31, 38, 41, 50, 55, 70, 72]. The main focus of seven articles lies on the perception of the visualizations of agents. Buck et al. [6] evaluate only self-avatars, Zibrek et al. [72] evaluate both the avatar of another person together with an agent. Two articles evaluate both self-avatars and agents, [21, 38]. Also, Latoschik et al. [31] evaluate self-perception in non-verbal dyadic communication using self-avatars together with the perception of the avatar of another person. Three studies contained a conversational task with an agent [47, 50, 55]. Three experiments involved a stop-distance task to investigate proxemics [6, 21, 72]. The other studies involved more passive activities like listening to an instructor [62], listening to a character in a story-based AR or VR experience [41, 71], or performing simple movements [31] to communicate with another person (non-verbal). Nine out of 12 studies focus on rendering style. Here, the conversation partner was often an agent or a real human that had to follow a rigid script. One study considered — next to a full-body visualization — a configuration which included no body at all [47]. Next to that, Wang et al. [62] focused on body part visibility of an abstract visualization, hands-only, hands+arms, and full-body. Similarly, Skarbez et al. [55] investigated feet-only and full-body. In the remaining papers, the rendering styles range from generic realistic (all studies) to personalized realistic [6, 38], translucent [47], robot [6, 21], abstract [21, 62], hybrid and real-video [64], stickman [31], cartoon [38, 41, 50, 70], and stylized [62, 70–72].

Our analysis revealed that realistic avatars and agents (generic and personalized) as well as full-body avatars and agents led to most of the significant benefits. For example, user experience was usually better with generic realistic avatars [47] or agents [21, 62]. Similarly, the uncanny valley was perceived as less of a problem with realistic full-body avatars of others or agents [31, 47]. Also, presence benefits from generic realistic full-body agents [55, 71, 72]. The same goes for social presence [72]. However, this benefit was not found for co-presence [21] — or only in combination with specific personality types [70]. Here, Zibrek et al. [70] found that the personality of an agent influences its perceived appeal. As an additional factor to consider, Latoschick et al. [31] report that the appearance of others affects self-embodiment (the more realistic the higher self-embodiment). We want to highlight the study of Schloss et al. [50] about children exploring social boundaries in VR and approaching virtual characters as intelligent agents. One study mentioned a fidelity-related aspects: In addition to that, Zibrek et al. [71] recommend to be careful with animations on realistic agents as a fidelity mismatch can lead to degraded emotional responses. Other variables, affinity [70, 71], emotional response [41, 71], or perceived realism [70, 71] were unaffected by the type of visualization.

Several studies report tangible benefits of generic realistic full-body avatars, for example with respect to usability and presence. Thus, a generic realistic full-body agent and avatar can be recommended for static/uni-directional communication scenarios together with a coherent agent animation fidelity and a matching agent's personality.

4.4 Game-like scenario

Eight papers were categorized as a game-like scenario [11, 14, 17, 34–37, 48]. Seven of them deal with VR and one with AR [48]. All authors investigate self-representation (self-avatar) except Freiwald et al. [11] who investigate the representation of other humans. Tasks in these studies are all game-like or playful. This includes a first person shooter [34], a game where a player has to touch as many spheres as possible within 2.5 minutes [36, 37], an escape room [35], a setting were a character needs to stab another character with a knife [48], a virtual snowball fight [11], a bow shooting game [17], and a game where participants have to find and collect objects [14]. Five studies focused on the influence of rendering style [11, 35–37, 48]. Three focused on the influence of visible body parts [14, 17, 34]. All studies except two include a full-body avatar. These two focus on the hands+arms [34, 37] or an avatar consisting of head+hands+arms+feet [34, 37]. Various other combinations like hands-only [14, 17], head+hands [34], hands+feet [17] hands+torso [14], and no avatar [14, 34] were also part of comparisons.

The studies used various rendering styles, including abstract [37, 48], cartoon [11], stick-man [35], stylized [48], and robot [35, 36]. All except two include a generic realistic version in the study [34, 37]. None used a personalized or scanned realistic avatars.

The various conditions in these experiments did not show significant differences with respect to variables such as game experience [35], simulator sickness [35, 36], enjoyment [11], or arousal [36]. The results of these papers suggest that the visible body parts or rendering style do also not impact embodiment [34], virtual body ownership [11, 37], or social presence [36]. The full-body robot avatar of Lugrin et al. [36] led to a stronger Proteus effect due to it being perceived as more "powerful" and less "fragile" (compared to a generic realistic human-like avatar). Task performance was mostly unaffected by style or visible body parts [34, 35, 35]. Only Gao et al. [14] report increased task performance with a full body avatar compared to hands only or hands+torso. Except for Goncalves et al. [17], presence was unaffected [11, 34, 36]. They report increased presence — surprisingly not with a full body avatar but when only hands were shown (generic realistic). They attribute that to a bad tracking model where the animation was only driven by the feet and hand. Thus, movements appeared unnatural.

Overall, it is interesting that engaging tasks that are embedded in a story do not show the benefits of more realistic full-body avatars (as it was found in previously, cf. Sect. 4.3). Also, outstanding is the study of Lugrin et al. [36] where participants shifted their behavior according to the visualization and took over the qualities of a robot (usually seen as strong and durable) — future research should investigate if this effect could also help in other domains (e.g., physical activity). Papers in this category often mention requiring appropriate tracking to realize movement congruence and to support body ownership.

4.5 Education/Training

We found five papers that involved tasks related to learning activities. Three evaluated agents [46, 49, 64], one evaluated the avatar of another person [15], and one evaluated both a self-avatar and an agent. [39] These papers include activities in VR such as a guided tour through a museum exhibition [46, 49], a virtual field trip [64], a recall and object finding task (spatial learning) [15], and giving lab instructions and explaining theoretical concepts to students [39]. The main goal of these papers was to investigate the instructor's representation [15, 39, 46, 49, 64] in learning activities in VR. Regarding appearance, upper-body [15] or full-body [39, 46, 49, 64] were used. The rendering styles included a generic realistic [39, 46, 49, 64], robot [15, 39, 49, 64]), and the video feed of the participant [15]. Three studies included a baseline condition without an avatar or agent and used only audio [15, 46, 49].

In the study of Rzayev et al. [49], the perceived human-likeness was higher when only audio was used and in George et al. [15], participants reported a higher presence with the generic realistic or video avatar. Among the other findings from this domain, Rzayev et al. [49] showed that the lack of a virtual body could favor learning outcomes as it could be perceived as less distracting. Additionally, George et al. [15] and Makransky et al. [39] showed that there is a negative effect of presence in learning performance, i.e., higher levels of presence led to lower performance when using a self-avatar or the avatar of another. However, Petersen et al. [46] stated that the influence of presence on learning outcomes may be dependent on the type of information to be learned. If the information is factual, the avatar instructor may affect learning negatively, but if the information is conceptual, an avatar may support learning. It is relevant to mention that the participants in the study of Petersen et al. [46] were expert VR users. Also, Makransky et al. [39] point to different outcomes for male and female learners, where male learners had better learning outcomes with a robot while females performed better with a female humanoid avatar. Interestingly, male learners experienced higher presence with the female humanoid avatar. The authors attribute this to male students paying less attention to the learning material with the female avatar.

In this task domain, the appearance of the avatars used was mostly humanoid or robot-like. Almost all studies evaluated the effect of presence. The findings suggest that an increase in presence has a negative effect on the learning outcome. This varies from previous studies which have pointed to a rather positive effect of presence in learning [103, 107]. Thus, there might be a wider range of related factors like gender [39] and the type of knowledge to be acquired [46] that need to be considered.

5 Discussion

5.1 High-level findings

Overall, the analyzed studies highlight that avatar visualization and body-part configuration have a significant impact on quantitative and qualitative measures in AR and VR environments. In the following, we summarize our findings.

(1) Articles that compare various body-part-visibility configurations almost always report significant advantages of a full-body avatar and agent compared to other configurations. The advantages range from task performance [13, 14, 39, 58, 61], communication behavior [13, 61], user experience [14, 16, 62], presence [13, 43], to social presence [20, 65]. We observed this phenomenon in all task domains except game-like scenarios and education, where the amount of visible body parts was not evaluated (education/training) or did not lead to significant improvements (game-like scenarions). However, workload [15], simulator sickness [34, 36], or agency [58] seem to be less affected by various body part configurations — the studies in our corpus did not report significant differences.

(2) A realistic rendering tends to outperform other rendering styles and often leads to significant improvements. This applies for generic, personalized, scanned realistic, 3D point clouds, and video avatars or agents. Most significant benefits were measured with these rendering styles across task domains except gaming. Looking at the individual variables, the statistically significant positive improvements have been shown with respect to embodiment [6, 10, 12, 43], body ownership [3, 18, 22, 24, 25, 30–33, 60, 68], presence [13, 25, 26, 44, 49, 52, 60, 68, 71], and social presence [4, 8, 15, 65, 72]. User experience also benefits from realistic representations [4, 10, 19, 47, 61, 63, 67] and so does task performance [6, 13, 19, 26, 27, 33, 39, 42, 43, 45, 59] as well as the quality of emotional responses [8, 43, 60, 70, 72]. A final benefit of realistic representations is that the uncanny valley seems to have less of a negative impact with these representations [31, 47, 49, 53, 54].

(3) The analysis of articles that investigated the perception and/or interaction with agents revealed that a generic realistic full-body agent performed best in almost all cases. While only 17 manuscripts compared various types of agents, this trend is rather strong, showing improvements with respect to agency, presence, realism, user experience, and less negative impacts of the uncanny valley [46, 47, 49, 55, 61, 69].

(4) Applications that require the user to be focused on the task itself, such as bow shooting or even gaming, do not benefit much from realistic looking avatars (cf. game-like scenarios Sect. 4.4 or education Sect. 4.5). Communication tasks seem to benefit to a great extent from realistic rendering styles (cf. Sect. 4.3.1 and Sect. 4.3.2). Point cloud avatars and videos could even perform better than generic realistic representations for dynamic communication where body language and subtle cues are important [8, 13, 67]. We believe that realistic visualizations [4, 13, 56, 61] are crucial in communication as they allow for non-verbal behavior, which are crucial to understand the information that is being transmitted.

(5) Configurations of visible body parts that show head+hands [20, 65], head+hands+arms+feet [34] or head+torso [65] did not lead to any significant improvements in the analyzed experiments, regardless of task domain. We observed similar results for hybrid visualizations (e.g., human face with robot body [64]).

(6) When comparing results of the AR and VR experiments, they tend to be very similar. However, none of the analyzed studies conducted a direct comparison between AR and VR. Only one study deals with asymmetric setups (e.g., one user in AR and one in VR). We believe that our high-level findings can be transferable from AR to VR or vice-versa. However, parameters such as field of view and color fastness are very different between VR-HMDs and AR-HMDs as well as AR-ready smartphones. Hence, we encourage further research in AR and asymmetric setups (cf. Cho et al. [80] or Karasmanoglu et al. [97]).

(7) Eleven articles explicitly compared various realistic avatar visualizations (point cloud, video avatars, generic/personalized/scanned realistic [4, 6, 8, 12, 13, 22, 25, 60, 63, 64, 67]). Almost all significant improvements measured by these articles happened in a condition where the user was either personalized realistic, a video avatar, or a point cloud (most often paired with a full-body representation) and not in the generic realistic condition or the scanned realistic condition. This was the case for body ownership, co-presence, likability, embodiment, user experience, and trust. Thus, while more realistic avatars come with a higher technical cost, their benefits are tangible.

5.2 Guidelines

The results derived from our analysis inspired the following guidelines:

1. If in doubt: use a generic realistic full-body or a generic realistic hands-only visualizations (cf. High-level findings (1), (2) and (3)).
2. The user should have agency over visible body parts (keywords: movement congruence or behavioral fidelity).
3. Realistic rendering styles are especially preferable for communication (cf. High-level findings (4)).
4. For engaging content, such as games or easy/routine tasks, realistic rendering is not essential, but control fidelity is (cf. Sect. 4.4).
5. Realistic hands are beneficial for hand interaction (cf. Sect. 4.2).
6. Limit the visual distraction of an avatar or agent during learning or educational applications as it can potentially impair the learning outcome (cf. Sect. 4.5).
7. Point cloud, video avatars, and personalized realistic avatars tend to perform best (cf. High-level findings (4)), but come with a higher technological cost.

Overall, our findings seem to agree with previous work from other domains (e.g., non-HMD experiments) that also recommend higher visual fidelity [81, 82, 123, 126]. Still, future work should make a structured effort to assess generalization for avatar and agent visualization, especially with respect to body part visibility.

Finally, we emphasize that the research community dedicated to investigating avatars and agents in AR and VR could benefit greatly from replicating seminal studies from 2005 – 2010 with up-to-date hardware. To enable that, we encourage researchers to provide descriptions of the tools used to create avatars (e.g., for scanning or personalizing), demo applications, videos, models, information on tracking systems/technologies, and information on how to evaluate the data. This will enable replication and verification and speed up research in the community by lowering development times and support open science efforts. Some platforms that could support these efforts are Zenodo, Open Science Foundation (OSF), or IEEE DataPort, together with existing tools like QuickVR [113], RocketBox [86], HeadBox [127], and MoveBox [85].

5.3 Future research opportunities

(1) **Facial animation and lip synchronization.** None of the investigated studies explore facial animation across various avatar rendering styles. However, facial expressions are a crucial aspect of non-verbal behavior, which can determine the tone of conversations and interactions in collaborative virtual environments. As modern devices enable tracking movements of the face, e.g., HTC Vive Facial tracker (https://www.vive.com/us/accessory/facial-tracker/) and many AR/VR-HMDs include eye-tracking, we encourage future exploration and use of facial animations in avatars. Additionally, the implementation of machine learning-based 3D pose estimation could be a way to enhance avatar animation [109, 114]. Closely related to advances in technology for facial animation and lip synchronization is behavioral realism, which we excluded from our corpus (see Sect. 2.5). However, behavioral realism for avatars in AR/VR is an area that could be further explored, especially beyond social virtual contexts — potentially in a dedicated review.

(2) **Customization and inclusion.** Most studies rely on 3D models of generic avatars which have rarely been customized. Investigating the effect of customization while including aspects like disabilities, skin color, and body shape can expand the understanding of different types of behaviors and self-presentation choices in virtual environments. Self-disclosure of people with disabilities within virtual environments is an area that could be further explored, e.g., Zhang et al. [131] present one of the few studies that explores avatar diversity. Their findings show that people with disabilities can reflect disabilities selectively depending on the disability with stronger personal attachment or to emphasize their capabilities and independence. In addition to that, we believe that voice recognition and customization could be further explored, especially in terms of voice personalization of self-avatars.

(3) **Realistic human-avatar interaction.** Research has taken a closer look at the interaction between humans and realistic agents for providing instructions in maintenance and repair [83], interpersonal motor coordination [78], or communication (cf. Sect. 4.3). These domains could benefit from haptic feedback. Only few of our reviewed papers from the hand interaction category used different types of haptic feedback [1, 19, 26, 30, 53]. Future work should investigate how haptic feedback influences and interacts with avatars and agents.

(4) **AR vs. VR and asymmetric setups.** In our corpus, only 6 papers evaluated or used avatars or agents in AR. We believe that further studies should investigate if results from VR studies can also be obtained in AR environments to better understand the differences and effects of an entirely computer-generated setting compared to one that features real surroundings. For instance, investigating the role of social presence and co-presence in AR when using different rendering styles, e.g., realistic vs. abstract representations in real-world scenarios such as meetings or classrooms. Similarly, the ongoing democratization of AR and VR and multi-platform setups motivate further research in asymmetric setups.

(5) **Longitudinal avatar/agent use.** We believe that future research should consider long-term studies. In spite of the challenges of longitudinal research, e.g., participant's dropout, constructs becoming invalid, high costs [98], it is crucial to better understand how the use of avatars evolves over time. Especially since factors such as technology exposure or novelty have proven to influence how people interact in VR [75]. Khojasteh & Won [101] presented one of the few studies that evaluated how interaction evolved over time, with findings pointing to new ways of communication in VR. This could potentially shed light on research directions for avatar and agent design and confirm if the described positive (and negative) outcomes hold true, even after long-term usage.

(6) **The effect of confounding variables.** Our analysis revealed that multiple studies point to the influence of gender when evaluating presence [39, 52], proxemics [21], and agency [29]. To the best of our knowledge, the effect of gender has mostly been evaluated in learning settings. We believe that future studies should consider the effect of gender when evaluating VR experiences in a wider range of scenarios and tasks, e.g., communication and physical activities.

In addition, some papers in the reviewed corpus investigate the "Proteus effect" [130]. Lugrin et al. [36] showed that a robot avatar was perceived as more powerful and less fragile compared to humanoid avatars in a game. It was shown that the "Proteus effect" can provide several tangible consequences in AR and VR [102, 119]. However, we think that the phenomenon is not sufficiently explored yet and should receive additional attention in future research. Age is another variable that may affect the perception and interaction in VR experiences. Most participants in the studies are between 7 and 36 years old, leaving aside people from other age groups, especially people over 40. Thus, the behavior and preferences of people outside this age range are yet to be explored. This are, for example, changing socio-political views and gender-conforming behavior, adopted self-perception, and motor skills [124]. Actively leveraging these influences can foster cooperation, collaboration, and performance via avatars. By that, it is a promising research direction and could help to exploit the benefits of social AR and VR applications for a diverse set of users.

(7) **Subjective vs. objective measures.** Our analysis showed that the majority of studies relied on post-hoc questionnaires which are prone to a plethora of response biases [128]. Thus, we encourage future research to integrate physiological and behavioral measures to mitigate these biases (e.g., gaze behavior vs. NASA TLX for workload [105, 106, 116]). Here, it is also crucial to develop specific metrics for important constructs such as presence, embodiment, or body ownership or to validate existing ones (cf. Halbig & Latoschik [89]). Additionally, few studies actually compare their results with real scenarios and are therefore limited in their external validity.

6 Conclusion

In this article, we provide a systematic review of 72 studies, published between 2015 and 2022, that compare various types of avatars and agents in AR and VR, using the PRISMA method. We focus on research that uses HMDs, features at least two types of avatars or agents, and includes a user study. The analysis of these studies resulted in five main task domains that were derived from the tasks that users had to perform during the experiments. The domains include physical activities (walking, non-walking), hand interaction (direct, indirect), communication (uni-, bi-directional), game-like activities, and education/training. Our findings show that full-body visualizations usually outperform partial body configurations. Also, in most studies, a realistic rendering yields better results than other rendering styles. Both findings hold true, regardless of avatars and agents, and when comparing AR and VR. However, we found that the task domain influences avatar selection as for example more task-focused activities (e.g., gaming or education) do not benefit much from highly realistic avatars or are in some cases even hampered by them. Finally, we point out that, while generic realistic visualizations work well, personalized avatars that resemble the actual user usually perform even better. Based on our findings, we present general guidelines for the usage and evaluation of avatars and agents in research and industry. Moreover, we identify important future research directions, namely facial animation and lip synchronization, avatar customization and inclusion, AR vs. VR and asymmetric setups, longitudinal avatar/agent use, the effect of confounding variables, and objective vs. subjective measures for evaluation. We are convinced that, in a time where AR and VR are pushing into the consumer market, our findings can provide a basis for research and encourage further development of applications that include avatars and agents. Our data including supporting figures and a BibTex file is available at `https://doi.org/10.5281/zenodo.7525054`.

Acknowledgments

This work has been partially funded through the CYTEMEX project (Free State of Thuringia, Germany, FKZ: 2018-FGI-0019), the MULTI-PARTIES project (Federal Ministry of Education and Research, Germany), and the CO-HUMANICS project (Carl Zeiss Foundation).

Reviewed Articles

[1] L. Afonso, P. Dias, C. Ferreira, and B. S. Santos. Effect of hand-avatar in a selection task using a tablet as input device in an immersive virtual environment. In *3DUI*. IEEE Computer Society, 2017.

[2] C. R. Agnew. Feeling close to a crab-thing in virtual reality: Does avatar appearance always matter in forming meaningful connections? a case study. *Frontiers in Virtual Reality*, 3, 2022.

[3] F. Argelaguet, L. Hoyet, M. Trico, and A. Lécuyer. The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *VR*. IEEE Computer Society, 2016.

[4] S. A. Aseeri and V. Interrante. The influence of avatar representation on interpersonal communication in virtual social environments. *IEEE Trans. Vis. Comput. Graph.*, 27(5), 2021.

[5] B. Bodenheimer and Q. Fu. The effect of avatar model in stepping off a ledge in an immersive virtual environment. In *SAP*. ACM, 2015.

[6] L. E. Buck, S. Chakraborty, and B. Bodenheimer. The impact of embodiment and avatar sizing on personal space in immersive virtual environments. *IEEE Trans. Vis. Comput. Graph.*, 28(5), 2022.

[7] M. Butz, D. Hepperle, and M. Wölfel. Influence of Visual Appearance of Agents on Presence, Attractiveness, and Agency in Virtual Reality. In M. Wölfel, J. Bernhardt, and S. Thiel, editors, *ArtsIT, Interactivity and Game Creation*, 2022.

[8] S. Cho, S. Kim, J. Lee, J. Ahn, and J. Han. Effects of volumetric capture avatars on social presence in immersive virtual environments. In *VR*. IEEE, 2020.

[9] Y. Choi, J. Lee, and S. Lee. Effects of locomotion style and body visibility of a telepresence avatar. In *VR*. IEEE, 2020.

[10] D. Dewez, L. Hoyet, A. Lécuyer, and F. Argelaguet. Do you need another hand? investigating dual body representations during anisomorphic 3d manipulation. *IEEE Trans. Vis. Comput. Graph.*, 28(5), 2022.

[11] J. P. Freiwald, J. Schenke, N. Lehmann-Willenbrock, and F. Steinicke. Effects of avatar appearance and locomotion on co-presence in virtual reality collaborations. In *MuC*. ACM, 2021.

[12] R. Fribourg, F. Argelaguet, A. Lécuyer, and L. Hoyet. Avatar and sense of embodiment: Studying the relative preference between appearance, control and point of view. *IEEE Trans. Vis. Comput. Graph.*, 26(5), 2020.

[13] G. Gamelin, A. Chellali, S. Cheikh, A. Ricca, C. Dumas, and S. Otmane. Point-cloud avatars to improve spatial communication in immersive collaborative virtual environments. *Pers. Ubiquitous Comput.*, 25(3), 2021.

[14] B. Gao, J. Lee, H. Tu, W. Seong, and H. Kim. The effects of avatar visibility on behavioral response with or without mirror-visual feedback in virtual environments. In *VR Workshops*. IEEE, 2020.

[15] C. George, M. Spitzer, and H. Hussmann. Training in IVR: investigating the effect of instructor design on social presence and performance of the VR user. In *VRST*. ACM, 2018.

[16] D. Gochfeld, K. Benzing, K. Laibson, and A. Coulombe. Avatar selection for live performance in virtual reality: A case study. In *GEM*. IEEE, 2019.

[17] G. Gonçalves, M. Melo, L. Barbosa, J. Vasconcelos-Raposo, and M. Bessa. Evaluation of the impact of different levels of self-representation and body tracking on the sense of presence and embodiment in immersive VR. *Virtual Real.*, 26(1), 2022.

[18] G. Gorisse, O. Christmann, S. Houzangbe, and S. Richir. From robot to virtual doppelganger: Impact of visual fidelity of avatars controlled in third-person perspective on embodiment and behavior in immersive virtual environments. *Frontiers Robotics AI*, 6, 2019.

[19] J. Grubert, L. Witzani, E. Ofek, M. Pahud, M. Kranz, and P. O. Kristensson. Effects of hand representations for typing in virtual reality. In *VR*. IEEE Computer Society, 2018.

[20] P. Heidicker, E. Langbehn, and F. Steinicke. Influence of avatar appearance on presence in social VR. In *3DUI*. IEEE Computer Society, 2017.

[21] A. Huang, P. Knierim, F. Chiossi, L. L. Chuang, and R. Welsch. Proxemics for human-agent interaction in augmented reality. In *CHI*. ACM, 2022.

[22] D. Jo, K. Kim, G. F. Welch, W. Jeon, Y. Kim, K. Kim, and G. J. Kim. The impact of avatar-owner visual similarity on body ownership in immersive virtual reality. In *VRST*. ACM, 2017.

[23] T. Joy, E. Ugur, and I. Ayhan. Trick the body trick the mind: avatar representation affects the perception of available action possibilities in virtual reality. *Virtual Real.*, 26(2), 2022.

[24] S. Jung and C. E. Hughes. The effects of indirect real body cues of irrelevant parts on virtual body ownership and presence. In *ICAT-EGVE*. Eurographics Association, 2016.

[25] S. Jung, C. Sandor, P. J. Wisniewski, and C. E. Hughes. Realme: the influence of body and hand representations on body ownership and presence. In *SUI*. ACM, 2017.

[26] P. Knierim, V. Schwind, A. M. Feit, F. Nieuwenhuizen, and N. Henze. Physical keyboards in virtual reality: Analysis of typing performance and effects of avatar hands. In *CHI*. ACM, 2018.

[27] M. Kocur, S. Graf, and V. Schwind. The impact of missing fingers in virtual reality. In *VRST*. ACM, 2020.

[28] R. Kondo, M. Sugimoto, K. Minamizawa, T. Hoshi, M. Inami, and M. Kitazaki. Illusory body ownership of an invisible body interpolated between virtual hands and feet via visual-motor synchronicity. *Scientific reports*, 8(1):1–8, 2018.

[29] C. Krogmeier and C. Mousas. Eye fixations and electrodermal activity during low-budget virtual reality embodiment. *Comput. Animat. Virtual Worlds*, 31(4-5), 2020.

[30] B. N. Krom, M. Catoire, A. Toet, R. J. E. V. Dijk, and J. B. F. van Erp. Effects of likeness and synchronicity on the ownership illusion over a moving virtual robotic arm and hand. In *WHC*. IEEE, 2019.

[31] M. E. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch. The effect of avatar realism in immersive social virtual realities. In *VRST*. ACM, 2017.

[32] L. Lin and S. Jörg. Need a hand?: how appearance affects the virtual hand illusion. In *SAP*. ACM, 2016.

[33] C. Lougiakis, A. Katifori, M. Roussou, and I. P. Ioannidis. Effects of virtual hand representation on interaction and embodiment in hmd-based virtual environments using controllers. In *VR*. IEEE, 2020.

[34] J. Lugrin, M. Ertl, P. Krop, R. Klupfel, S. Stierstorfer, B. Weisz, M. Ruck, J. Schmitt, N. Schmidt, and M. E. Latoschik. Any "body" there? avatar visibility effects in a virtual reality game. In *VR*. IEEE Computer Society, 2018.

[35] J. Lugrin, J. Latt, and M. E. Latoschik. Avatar anthropomorphism and illusion of body ownership in VR. In *VR*. IEEE Computer Society, 2015.

[36] J. Lugrin, I. Polyschev, D. Roth, and M. E. Latoschik. Avatar anthropomorphism and acrophobia. In *VRST*. ACM, 2016.

[37] J. Lugrin, M. Wiedemann, D. Bieberstein, and M. E. Latoschik. Influence of avatar realism on stressful situation in VR. In *VR*. IEEE Computer Society, 2015.

[38] F. Ma and X. Pan. Visual fidelity effects on expressive self-avatar in virtual

reality: First impressions matter. In *VR*. IEEE, 2022.
[39] G. Makransky, P. Wismer, and R. E. Mayer. A gender matching effect in learning with pedagogical agents in an immersive virtual reality science simulation. *J. Comput. Assist. Learn.*, 35(3), 2019.
[40] Y. Matsuda, J. Nakamura, T. Amemiya, Y. Ikei, and M. Kitazaki. Enhancing virtual walking sensation using self-avatar in first-person perspective and foot vibrations. *Frontiers Virtual Real.*, 2, 2021.
[41] C. Mousas, D. Anastasiou, and O. Spantidi. The effects of appearance and motion of virtual characters on emotional reactivity. *Comput. Hum. Behav.*, 86, 2018.
[42] N. Ogawa, T. Narumi, and M. Hirose. Virtual hand realism affects object size perception in body-based scaling. In *VR*. IEEE, 2019.
[43] N. Ogawa, T. Narumi, and M. Hirose. Effect of avatar appearance on detection thresholds for remapped hand movements. *IEEE Trans. Vis. Comput. Graph.*, 27(7), 2021.
[44] N. Ogawa, T. Narumi, H. Kuzuoka, and M. Hirose. Do you feel like passing through walls?: Effect of self-avatar appearance on facilitating realistic behavior in virtual environments. In *CHI*. ACM, 2020.
[45] Y. Pan and A. Steed. The impact of self-avatars on trust and collaboration in shared virtual environments. *PLOS ONE*, 12(12):e0189078, December 2017.
[46] G. B. Petersen, A. Mottelson, and G. Makransky. Pedagogical agents in educational VR: an in the wild study. In *CHI*. ACM, 2021.
[47] J. Reinhardt, L. Hillen, and K. Wolf. Embedding conversational agents into AR: invisible or with a realistic human body? In *Tangible and Embedded Interaction*. ACM, 2020.
[48] N. Rosa, J. van Bommel, W. Hürst, T. Nijboer, R. C. Veltkamp, and P. J. Werkhoven. Embodying an extra virtual body in augmented reality. In *VR*. IEEE, 2019.
[49] R. Rzayev, G. Karaman, K. Wolf, N. Henze, and V. Schwind. The effect of presence and appearance of guides in virtual reality exhibitions. In *MuC*. GI / ACM, 2019.
[50] I. Schloss, J. O. Bailey, and S. Tripathi. "i'm in his belly!": Children's responses to different types of characters in virtual reality. In *IDC*. ACM, 2021.
[51] V. Schwind, P. Knierim, L. L. Chuang, and N. Henze. "where's pinky?": The effects of a reduced number of fingers in virtual reality. In *CHI PLAY*. ACM, 2017.
[52] V. Schwind, P. Knierim, C. Tasci, P. Franczak, N. Haas, and N. Henze. "these are not my hands!": Effect of gender on the perception of avatar hands in virtual reality. In *CHI*. ACM, 2017.
[53] V. Schwind, L. Lin, M. D. Luca, S. Jörg, and J. Hillis. Touch with foreign hands: the effect of virtual hand appearance on visual-haptic integration. In *SAP*. ACM, 2018.
[54] V. Schwind, S. Mayer, A. Comeau-Vermeersch, R. Schweigert, and N. Henze. Up to the finger tip: The effect of avatars on mid-air pointing accuracy in virtual reality. In *CHI PLAY*. ACM, 2018.
[55] R. Skarbez, S. Neyret, F. P. B. Jr., M. Slater, and M. C. Whitton. A psychophysical experiment regarding components of the plausibility illusion. *IEEE Trans. Vis. Comput. Graph.*, 23(4), 2017.
[56] H. J. Smith and M. Neff. Communication behavior in embodied virtual reality. In *CHI*. ACM, 2018.
[57] Y. Sun, O. Shaikh, and A. S. Won. Nonverbal synchrony in virtual reality. *PLOS ONE*, 14(9), September 2019.
[58] T. Q. Tran, H. Shin, W. Stuerzlinger, and J. Han. Effects of virtual arm representations on interaction in virtual environments. In *VRST*. ACM, 2017.
[59] D. Valkov, J. Martens, and K. H. Hinrichs. Evaluation of the effect of a virtual avatar's representation on distance perception in immersive virtual environments. In *VR*. IEEE Computer Society, 2016.
[60] T. Waltemate, D. Gall, D. Roth, M. Botsch, and M. E. Latoschik. The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE Trans. Vis. Comput. Graph.*, 24(4), 2018.
[61] I. Wang, J. Smith, and J. Ruiz. Exploring virtual agents for augmented reality. In *CHI*. ACM, 2019.
[62] T. Wang, Y. Sato, M. Otsuki, H. Kuzuoka, and Y. Suzuki. Effect of full body avatar in augmented reality remote collaboration. In *VR*. IEEE, 2019.
[63] M. Wirth, S. Gradl, G. Prosinger, F. Kluge, D. Roth, and B. M. Eskofier. The impact of avatar appearance, perspective and context on gait variability and user experience in virtual reality. In *VR*. IEEE, 2021.
[64] J. W. Woodworth, N. G. Lipari, and C. W. Borst. Evaluating teacher avatar appearances in educational VR. In *VR*. IEEE, 2019.
[65] B. Yoon, H. Kim, G. A. Lee, M. Billinghurst, and W. Woo. The effect of avatar appearance on social presence in an augmented reality remote collaboration. In *VR*. IEEE, 2019.
[66] M. K. Young, J. J. Rieser, and B. Bodenheimer. Dyadic interactions with avatars in immersive virtual environments: high fiving. In *SAP*. ACM, 2015.
[67] K. Yu, G. Gorbachev, U. Eck, F. Pankratz, N. Navab, and D. Roth. Avatars for teleconsultation: Effects of avatar embodiment techniques on user perception in 3d asymmetric telepresence. *IEEE Trans. Vis. Comput. Graph.*, 27(11), 2021.
[68] J. Zhang, M. Huang, L. Zhao, R. Yang, H. Liang, J. Han, L. Wang, and W. Sun. Influence of hand representation design on presence and embodiment in virtual environment. In *ISCID*. IEEE, 2020.
[69] K. Zibrek, E. Kokkinara, and R. McDonnell. 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 *SAP*. ACM, 2017.
[70] K. Zibrek, E. Kokkinara, and R. McDonnell. The effect of realistic appearance of virtual characters in immersive environments - does the character's personality play a role? *IEEE Trans. Vis. Comput. Graph.*, 24(4), 2018.
[71] K. Zibrek, S. Martin, and R. McDonnell. Is photorealism important for perception of expressive virtual humans in virtual reality? *ACM Trans. Appl. Percept.*, 16(3), 2019.
[72] K. Zibrek and R. McDonnell. Social presence and place illusion are affected by photorealism in embodied VR. In *MIG*. ACM, 2019.

Other References

[73] J. N. Bailenson and J. Blascovich. Avatars. In *Encyclopedia of human-computer interaction, Berkshire Publishing Group*, 2004.
[74] J. N. Bailenson, K. R. Swinth, C. L. Hoyt, S. Persky, A. Dimov, and J. Blascovich. The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence Teleoperators Virtual Environ.*, 14(4), 2005.
[75] J. N. Bailenson, N. Yee, D. Merget, and R. Schroeder. The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. *Presence Teleoperators Virtual Environ.*, 15(4), 2006.
[76] D. Banakou, S. Kishore, and M. Slater. Virtually Being Einstein Results in an Improvement in Cognitive Task Performance and a Decrease in Age Bias. *Frontiers in Psychology*, 9, 2018.
[77] F. Biocca and C. Harms. Defining and measuring social presence. In *In F.R. Gouveia, & F. Biocca (Eds). Proceedings of the 5th Intern. Workshop on Presence*, 2002.
[78] S. Boukarras, D. G. Özkan, V. Era, Q. Moreau, G. Tieri, and M. Candidi. Midfrontal theta transcranial alternating current stimulation facilitates motor coordination in dyadic human-avatar interactions. *Journal of cognitive neuroscience*, 34(5):897–915, 2022.
[79] J. Brooke. SUS: A quick and dirty usability scale. *Usability Eval. Ind.*, 189, Nov 1995.
[80] Y. Cho, J. Kang, J. Jeon, J. Park, M. Kim, and J. Kim. X-person asymmetric interaction in virtual and augmented realities. *Comput. Animat. Virtual Worlds*, 32(5), 2021.
[81] S. Daher, J. Hochreiter, R. Schubert, L. Gonzalez, J. Cendan, M. Anderson, D. A. Diaz, and G. F. Welch. The Physical-Virtual Patient Simulator: A Physical Human Form With Virtual Appearance and Behavior. *Simulation in Healthcare*, 15(2):115, Apr. 2020.
[82] Z. Dai and K. F. MacDorman. The doctor's digital double: How warmth, competence, and animation promote adherence intention. *PeerJ Computer Science*, 4:e168, Nov. 2018.
[83] S. D'Onofrio, R. Vertucci, S. Ricciardi, and M. de Nino. Augmented reality for maintenance and repair. In A. Y. C. Nee and S. K. Ong, editors, *Springer Handbook of Augmented Reality*, Springer Handbooks, pages 597–616. Springer International Publishing, Cham, 2023.
[84] A. Genay, A. Lécuyer, and M. Hachet. Being an avatar "for real": A survey on virtual embodiment in augmented reality. *IEEE Trans. Vis. Comput. Graph.*, 28(12), 2022.
[85] M. González-Franco, Z. Egan, M. Peachey, A. Antley, T. Randhavane, P. Panda, Y. Zhang, C. Y. Wang, D. F. Reilly, T. C. Peck, A. S. Won, A. Steed, and E. Ofek. Movebox: Democratizing mocap for the microsoft rocketbox avatar library. In *AIVR*. IEEE, 2020.
[86] M. González-Franco, E. Ofek, Y. Pan, A. Antley, A. Steed, B. Spanlang, A. Maselli, D. Banakou, N. Pelechano, S. Orts-Escolano, V. Orvalho,

L. Trutoiu, M. Wojcik, M. V. Sanchez-Vives, J. N. Bailenson, M. Slater, and J. Lanier. The rocketbox library and the utility of freely available rigged avatars. *Frontiers Virtual Real.*, 1, 2020.

[87] M. González-Franco, A. Steed, S. Hoogendyk, and E. Ofek. Using facial animation to increase the enfacement illusion and avatar self-identification. *IEEE Trans. Vis. Comput. Graph.*, 26(5), 2020.

[88] R. Green, P. H. Delfabbro, and D. L. King. Avatar- and self-related processes and problematic gaming: A systematic review. *Addictive Behaviors*, 108:106461, 2020.

[89] A. Halbig and M. E. Latoschik. A systematic review of physiological measurements, factors, methods, and applications in virtual reality. *Frontiers Virtual Real.*, 2, 2021.

[90] B. S. Hantono, L. E. Nugroho, and P. I. Santosa. Review of augmented reality agent in education. In *2016 6th International Annual Engineering Seminar (InAES)*, pages 150–153, August 2016.

[91] S. G. Hart. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, October 2006.

[92] B. S. Hasler, B. Spanlang, and M. Slater. Virtual race transformation reverses racial in-group bias. *PLOS ONE*, 12(4):e0174965, April 2017.

[93] M. Hassenzahl, M. Burmester, and F. Koller. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In G. Szwillus and J. Ziegler, editors, *Mensch & Computer 2003: Interaktion in Bewegung*, Berichte des German Chapter of the ACM, pages 187–196. Vieweg+Teubner Verlag, Wiesbaden, 2003.

[94] C.-C. Ho and K. F. MacDorman. Measuring the Uncanny Valley Effect. *International Journal of Social Robotics*, 9(1):129–139, January 2017.

[95] I. Hudson and J. Hurter. Avatar types matter: Review of avatar literature for performance purposes. In *HCI (10)*, volume 9740 of *Lecture Notes in Computer Science*. Springer, 2016.

[96] igroup.org. Igroup presence questionnaire (IPQ) overview — igroup.org – project consortium. http://www.igroup.org/pq/ipq/index.php, 2016.

[97] S. Karaosmanoglu, K. Rogers, D. Wolf, E. Rukzio, F. Steinicke, and L. E. Nacke. Feels like team spirit. In *CHI*. ACM, 2021.

[98] E. Karapanos, J. Gerken, J. Kjeldskov, and M. B. Skov. Introduction to "advances in longitudinal hci research". In E. Karapanos, J. Gerken, J. Kjeldskov, and M. B. Skov, editors, *Advances in Longitudinal HCI Research*, Human–Computer Interaction Series, pages 1–7. Springer International Publishing, Cham, 2021.

[99] J. Kätsyri, K. Förger, M. Mäkäräinen, and T. Takala. A review of empirical evidence on different uncanny valley hypotheses: Support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in Psychology*, 6, 2015.

[100] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, 3(3):203–220, July 1993.

[101] N. Khojasteh and A. S. Won. Working together on diverse tasks: A longitudinal study on individual workload, presence and emotional recognition in collaborative virtual environments. *Frontiers Virtual Real.*, 2, 2021.

[102] M. Kocur. *Utilizing the Proteus Effect to Improve Performance Using Avatars in Virtual Reality*. PhD thesis, Regensburg University, GER, 2022.

[103] A. M. Kontogeorgiou, J. Bellou, and T. A. Mikropoulos. Being inside the quantum atom. *PsychNology J.*, 6(1), 2008.

[104] M. Korban and X. Li. A survey on applications of digital human avatars toward virtual co-presence. *CoRR*, abs/2201.04168, 2022.

[105] F. Li, G. Xu, and S. Feng. Eye tracking analytics for mental states assessment - A review. In *SMC*. IEEE, 2021.

[106] G. Marquart, C. Cabrall, and J. de Winter. Review of Eye-related Measures of Drivers' Mental Workload. *Procedia Manufacturing*, 3:2854–2861, January 2015.

[107] R. E. Mayer. Principles of multimedia learning based on social cues : Personalization, voice, and image principles. In R. E. Mayer, editor, *The Cambridge handbook of multimedia learning*, pages 201–212. University of Cambridge, New York, 2005.

[108] R. McDonnell. Appealing virtual humans. In *MIG*, volume 7660 of *Lecture Notes in Computer Science*. Springer, 2012.

[109] D. Mehta, S. Sridhar, O. Sotnychenko, H. Rhodin, M. Shafiei, H. Seidel, W. Xu, D. Casas, and C. Theobalt. Vnect: real-time 3d human pose estimation with a single RGB camera. *ACM Trans. Graph.*, 36(4), 2017.

[110] N. Norouzi, G. Bruder, B. Belna, S. Mutter, D. Turgut, and G. Welch. A systematic review of the convergence of augmented reality, intelligent virtual agents, and the internet of things. In F. Al-Turjman, editor, *Art. Intel. in IoT*, Transa. on Comp. Science and Comp. Intel., pages 1–24. Springer International, Cham, 2019.

[111] C. S. Oh, J. N. Bailenson, and G. F. Welch. A systematic review of social presence: Definition, antecedents, and implications. *Frontiers Robotics AI*, 5, 2018.

[112] C. Oh Kruzic, D. Kruzic, F. Herrera, and J. Bailenson. Facial expressions contribute more than body movements to conversational outcomes in avatar-mediated virtual environments. *Scientific Reports*, 10(1):20626, Nov 2020.

[113] R. Oliva, A. Beacco, X. Navarro, and M. Slater. QuickVR: A standard library for virtual embodiment in unity. *Frontiers in Virtual Reality*, 3, 2022.

[114] P. Panteleris, I. Oikonomidis, and A. A. Argyros. Using a single RGB frame for real time 3d hand pose estimation in the wild. In *WACV*. IEEE Computer Society, 2018.

[115] T. C. Peck and M. González-Franco. Avatar embodiment. A standardized questionnaire. *Frontiers Virtual Real.*, 1, 2020.

[116] S. Peißl, C. D. Wickens, and R. Baruah. Eye-Tracking Measures in Aviation: A Selective Literature Review. *The International Journal of Aerospace Psychology*, 28(3-4):98–112, October 2018.

[117] C. Pickering, J. Grignon, R. Steven, D. Guitart, and J. Byrne. Publishing not perishing: How research students transition from novice to knowledgeable using systematic quantitative literature reviews. *Studies in Higher Education*, 40(10), Nov 2015.

[118] I. V. Piryankova, H. Y. Wong, S. A. Linkenauger, C. Stinson, M. R. Longo, H. H. Bülthoff, and B. J. Mohler. Owning an Overweight or Underweight Body: Distinguishing the Physical, Experienced and Virtual Body. *PLOS ONE*, 9(8):e103428, August 2014.

[119] A. S. Praetorius and D. Görlich. How avatars influence user behavior: A review on the proteus effect in virtual environments and video games. In *FDG*. ACM, 2020.

[120] J. M. Rigby, D. P. Brumby, S. J. J. Gould, and A. L. Cox. Development of a questionnaire to measure immersion in video media: The film IEQ. In *TVX*. ACM, 2019.

[121] D. Roth, J. Lugrin, M. E. Latoschik, and S. Huber. Alpha IVBO - construction of a scale to measure the illusion of virtual body ownership. In *CHI Extended Abstracts*. ACM, 2017.

[122] R. B. Rubin, P. Palmgreen, and H. E. Sypher. Interpersonal Communication Satisfaction Inventory. In *Communication Research Measures*. Routledge, 2004.

[123] J. Stuart, K. Aul, A. Stephen, M. D. Bumbach, and B. Lok. The Effect of Virtual Human Rendering Style on User Perceptions of Visual Cues. *Frontiers in Virtual Reality*, 3, 2022.

[124] K. Szolin, D. J. Kuss, F. M. Nuyens, and M. D. Griffiths. Exploring the user-avatar relationship in videogames: A systematic review of the Proteus effect. *Human–Computer Interaction*, 0(0):1–26, August 2022.

[125] M. Usoh, E. Catena, S. Arman, and M. Slater. Using presence questionnaires in reality. *Presence Teleoperators Virtual Environ.*, 9(5), 2000.

[126] M. Volonte, A. T. Duchowski, and S. V. Babu. Effects of a Virtual Human Appearance Fidelity Continuum on Visual Attention in Virtual Reality. In *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, IVA '19, pages 141–147, New York, NY, USA, July 2019. Association for Computing Machinery.

[127] M. Volonte, E. Ofek, K. Jakubzak, S. Bruner, and M. González-Franco. Headbox: A facial blendshape animation toolkit for the microsoft rocketbox library. In *VR Workshops*. IEEE, 2022.

[128] P. A. Williams, J. L. Jenkins, J. S. Valacich, and M. D. Byrd. Measuring actual behaviors in HCI research - A call to action and an example. *AIS Trans. Hum. Comput. Interact.*, 9(4), 2017.

[129] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence Teleoperators Virtual Environ.*, 7(3), 1998.

[130] N. Yee and J. Bailenson. The Proteus Effect: The Effect of Transformed Self-Representation on Behavior. *Human Communication Research*, 33(3):271–290, July 2007.

[131] K. Zhang, E. Deldari, Z. Lu, Y. Yao, and Y. Zhao. "it's just part of me: " understanding avatar diversity and self-presentation of people with disabilities in social virtual reality. In *ASSETS*. ACM, 2022.