Avatar quality: A study on presence and user preference

Cristina Gasch  
*Universitat Jaume I*  
Spain  
cgasch@uji.es

Alireza Javanmardi  
German Research Center for Artificial Intelligence (DFKI)  
Germany

Azucena Garcia-Palacios  
*Universitat Jaume I*  
Spain

Alain Pagani  
German Research Center for Artificial Intelligence (DFKI)  
Germany

Abstract—With the increasing demand for remote work, virtual reality technologies are increasingly being considered as an option. However, the cybersickness they can induce may hinder their usage. A factor that can help mitigate this issue is allowing a user representation to interact in the world more naturally, thereby increasing the sense of presence. The aim of this study is to obtain a comparative analysis of the user’s sense of presence when using different avatars in a work and educational context. Additionally, it aims to analyze user acceptability and preferences for using them. The study involved 42 users ranging from 22 to 62 years of age, all of whom had prior experience with technologies for online meetings or virtual reality. The results indicate that hyper-realistic avatars generate a greater sense of presence, and users prefer them in work and educational settings. Furthermore, the results reveal a high intention to use them in these contexts.

Index Terms—Avatar, Virtual Reality, User Presence

I. INTRODUCTION

In recent years, the importance of seeking a better balance between work and personal life has increased. One of the changes driven by this need and technological advancements has been remote work. Therefore, tools that enable employees to perform their tasks from home are becoming more common [1]. However, this mode of work is not without challenges, such as distractions, the absence of a suitable workspace, difficulty in maintaining a schedule, and separating work from everyday life, among others [2].

To mitigate these problems and enable high-quality telecommuting, some studies have considered that virtual reality not only enhances tools for remote work, such as understanding the environment to interact with it [3], providing access to materials not normally available [4], and creating training scenarios that are challenging in a physical setting, as seen in medicine [5], [6]. It also shows promising perspectives for increasing immersion and the sense of presence [7], [8], improving efficiency, reducing distractions from the physical world, and enabling the implementation of better work-life balance strategies [2].

However, there are still numerous challenges for the everyday use of virtual reality in office-like tasks. These challenges include virtual motion sickness, which encompasses symptoms such as visual fatigue, muscular fatigue, acute stress, and mental overload caused by the use of virtual reality hardware [9], [10].

There are several casual factors that contribute to the onset of virtual motion sickness, such as demographic aspects [11], the illusion of motion [12], [13], occlusion of peripheral vision [14], or visualization and rendering modes [15]. However, Weech et al. [16] suggest that certain factors, such as sensory adjustments and increased interaction, reduce virtual motion sickness and enhance the user’s sense of presence in the environment. Mayor et al. [17] support this statement by comparing different types of interaction, with natural movement generating greater presence and less cybersickness. Therefore, making the user feel transported into the virtual environment and allowing them to interact naturally will lead to a reduction of these symptoms.

To achieve this increase in interaction with the environment, it is often necessary to include a representation of the user within the environment, allowing them to embody a virtual avatar [18]. This will not only enhance the sense of presence but also enable interaction with other users, for example, in a virtual meeting.

However, incorporating user avatars in a virtual environment can lead to potential problems if not done correctly, affecting the overall experience and giving rise to negative changes in attitude, behavior, and cognition [19]. On one hand, using avatars that closely resemble reality can trigger feelings of insecurity while using the tool [20]. On the other hand, allowing avatar customization may lead to altered behavior [21] or the development of psychological conditions like body dysmorphia [22], [23]. Using avatars that approach reality but do not cross into hyper-realistic territory can help mitigate some of the self-esteem-related problems and avoid insecurity [24].

However, this reduction in avatar realism could result in a decrease in user presence. Nevertheless, this reduction may not be significant, considering that factors such as tracking level, stereoscopy, and the user’s field of vision have a greater impact on user presence compared to image quality, resolution, and sound [7].

To explore whether differences in avatars affect the sense of presence, acceptability, and feasibility, a study has been conducted. This study analyzed three types of avatars: hyper-realistic, non-realistic, and the use of hyper-realistic avatars.
belonging to others. The goal is to understand the benefits of each type of avatar and gather user preferences in different contexts in an initial study to establish a starting point for further exploration in the future.

II. Study Design

The study consisted of a classic usability test [25] conducted in a laboratory setting, simulating an online meeting. The session was divided into three tests, with each test using a different type of avatar. During each test, users familiarized themselves with the system and followed the researcher’s instructions to gather quantitative data. The study was conducted individually for each participant.

A. Participants

The inclusion criteria were: Age between 18 and 65 years, previous experience with online meetings or virtual reality systems. Sufficient cognitive, auditory, and/or visual abilities to read, write, or engage in a conversation in the language used in the study. The exclusion criteria were: Null or very limited ICT skills and sensory disorders that make participation in the study very difficult (such as blindness or deafness).

The final sample consisted of a total of 42 participants, 22 men and 20 women. The participants’ ages ranged from 22 to 62 years, with a mean age of 32.4 years (SD = 9.3). The literacy level was high, with 71.5% having a university degree or higher, of which 16.7% held a doctorate. The study was conducted in two countries: 32 participants from Spain and 10 from Germany.

Regarding their prior experience with online meetings in various contexts, 69% use them frequently, and 31% have used them more than 10 times. Concerning their experience with virtual reality technology, 23% use it frequently, 61.9% have used it at least once, and 14.3% have never used it.

B. Materials

To simulate the use of an avatar with different appearances in an online meeting, we utilized the state-of-the-art deep learning-based face reenactment model FNNeVR [26]. This model can animate face images namely as source based on video frames called driving frames. We selected FNeVR for our study as it provides high-fidelity face generation and real-time rendering speed. In addition to the possibility of appearance alteration, FNeVR can also be used for video compression in a video conference system by only transmitting a sparse set of facial key points and re-synthesizing the face on the receiver side. This combination of video compression and alternate appearance capabilities offers new possibilities for immersive video communication experiences. The combination of these features makes recent face-animating models a promising tool for developing new and innovative video communication experiences.

This technology finds applications in two distinct scenarios: same-identity and cross-identity reenactment. In the first scenario, both the source image and the driving frames, belong to the same individual using the webcam. However, in the second scenario, these components can pertain to two different identities as shown in Fig. 1.

C. Variables and measuring instruments

The evaluation protocol consisted of three parts: User information collected before the task, including sociodemographic data and prior experience with online meetings and VR technology; Information about the level of presence collected during the tests; And the information gathered at the end of the study, which included user acceptability and preferences data.

1) User profile: To obtain the user profile, a series of data was collected before initiating the test, including: Demographic information (gender, age, and educational background); User’s experience with online meetings in different contexts (work, education, and leisure) and the user’s previous experience with virtual reality technologies in these same contexts.

2) Measures obtained at the end of each test: To obtain the measurements regarding the user’s opinion at the end of each test, a questionnaire based on a 5-point Likert scale that ranges from Strongly Disagree to Strongly Agree adapted from the System Usability Scale (SUS) was used [27]. To measure the user’s sense of presence in each test, a psychometric approach was used to evaluate the levels of embodiment towards an artificial body part, specifically the face. The latent variables of property, agency, and change were identified [28]. Property refers to perceiving the body as one’s own, as a source of sensations. Agency is linked to the sensation of control over one’s own actions. And change refers to the experience of the self-located in the position of our body. The questions were adapted from a validated test [29], with some modifications to focus solely on the face. These privacy-related questions were included in the
specific questionnaires for each test as they were considered more relevant in the context of each evaluation. The questions used in the tests are displayed in Table I.

### TABLE I
**USER PRESENCE QUESTIONS**

<table>
<thead>
<tr>
<th>Property</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>It felt like the virtual face was my face.</td>
<td>The virtual face felt like a human face.</td>
<td>I had the feeling that the virtual face belonged to me.</td>
</tr>
<tr>
<td>Agency</td>
<td>Q4</td>
<td>Q5</td>
<td>Q6</td>
</tr>
<tr>
<td></td>
<td>The movements of the virtual face seemed to be my own movements.</td>
<td>I enjoyed controlling the virtual face.</td>
<td>I have felt comfortable using the virtual face.</td>
</tr>
<tr>
<td>Change</td>
<td>Q9</td>
<td>Q10</td>
<td>Q11</td>
</tr>
<tr>
<td></td>
<td>I had the illusion of owning a different face from my own.</td>
<td>I felt the need to check if my face really still looked like what I had in mind.</td>
<td>I felt as if the form or appearance of my face had changed.</td>
</tr>
<tr>
<td>Privacy</td>
<td>Q12</td>
<td>Q13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I feel that the use of this type of avatar is an intrusion into my privacy.</td>
<td>I feel that this kind of avatar reveals private personal information without my consent.</td>
<td></td>
</tr>
</tbody>
</table>

3) **Measures obtained at the end of the study:** After completing all four tests with different avatar representations, the participants provided their opinions on the experiences. This final evaluation compared the four tests and was also based on the users’ opinions. This evaluation consists of two parts:

The first part is an acceptability test to determine perceived usefulness and intention to use. For this part, a distinction was made between usage contexts since perceived usefulness and intention to use may vary depending on them: Work, education, and leisure. The scale used is again the SUS scale, and the questions have been extracted and adapted from a validated acceptability questionnaire [31]–[33].

Finally, the second part aims to understand the user’s preferences regarding the avatar to be used in different situations. The questionnaire includes six avatar usage scenarios: Team meetings or with external participants; Education as a student or as a teacher; and leisure with strangers, like games, or with friends. For each situation, multiple responses can be selected: Hyper-realistic avatar, realistic avatar, non-realistic avatar, or an avatar that does not represent the user.

### D. Hardware

All users conducted the study individually on a computer with an NVIDIA RTX3060 graphics card through local access. This graphics card enables the execution of the software with minimal delay in the user’s image and movements.

### E. Procedure

The study for each user was conducted in the following phases:

1) **PRE-Assessment:** In this phase, the user was informed about the study and what they would be doing. If they agreed to participate, they signed the informed consent document.

2) **PRE-Questionnaire:** The user’s sociodemographic data and profile were collected in this phase.

3) **Study (first part):** Three tests were conducted in a single session, with each test using a different photograph. Fig. 2 shows an example of the three photographs used to simulate the avatars in the tests. Each user always performed the four tests in the same order to avoid any impact related to differences in the sequence. In each test, the following steps were carried out.

![Fig. 2. Left: Hyper-realistic; Center: Non-realistic; Right: Other person.](image)

a) Avatar description: The user was provided with an explanation of the type of avatar being used, which included three options:

- Hyper-realistic: A digitally created representation of a human being with an exceptionally high level of detail and realism.
- Non-realistic: A digital representation of a human being intentionally deviating from realistic appearance and features, adopting stylized or abstract styles.
- Another person: A hyper-realistic digital representation of another person of the same gender who is not the user.

b) Familiarization: The user became acquainted with the avatar by performing head movements and gestures for 15 to 30 seconds.

c) Interview: While the user kept their gaze on their avatar, a researcher positioned behind the user to avoid drawing their attention simulated an interview with the user, asking casual questions. The user had the option to decline to answer, and if they felt uncomfortable, they could request a change of question. The objective was to
encourage conversation, divert attention from the avatar, and observe its movements during responses for 4 to 5 minutes. A sample of how the users conducted the study can be seen in the photograph taken during the study in Fig. 3.

d) Questionnaire: While the researcher prepared for the next experiment, the user completed the corresponding questionnaire.

4) Study (second part): For the final test, the camera captured the interviewer’s movements using the hyper-realistic photograph of the user. This way, the user could see how another person was using their own avatar. The data collection process remained the same as in the first part, using the interview format and answering a fourth questionnaire.

5) POST-Study: Finally, the user completed the POST-Study questionnaire.

III. RESULTS
A. User’s sense of presence
At the end of each test, the users completed a questionnaire regarding the type of avatar used. The questions in this questionnaire assessed the sense of being present in the virtual environment. To do so, the questions were divided into the three properties on which the sense of presence depends: property, agency, and change.

In the sense of property, it can be seen in Fig. 4 how both hyper-realistic avatars felt more like human faces than the non-realistic one (Q2). However, it was the user’s hyper-realistic avatar that scored higher in the questions regarding facial ownership (Q1 and Q3).

For the Agency property, four questions were conducted. In Fig. 5, it is shown how the user’s avatars had higher scores in all questions. The hyper-realistic avatar slightly scored better in questions about avatar movement (Q4, Q7, and Q8). The non-realistic avatar scored higher in questions related to enjoyment (Q5) and comfort (Q6). In contrast, the avatar of another person did not score high in any question except for the comfortable sense.

Finally, the results for the change property can be seen in Fig. 6. In all three cases, the values were quite low (values less than 2.5). However, it can be noted that the non-realistic avatar produced the most pronounced sense of change. Interestingly, the hyper-realistic avatar made more users question whether their face had changed (Q10). It is also worth highlighting that the avatar not belonging to the user obtained the lowest scores in all three questions.

For the three factors, correlation with sociodemographic data and prior experience was analyzed using a Spearman correlation analysis. The results showed that there is no significant correlation with age, gender, education level, or prior experience with videoconferencing. However, there is a correlation with prior experience with virtual reality. Table II displays the obtained results, differentiating between types of
avatars: Hyper-realistic (H), non-realistic (N), and the avatar of another realistic person (O).

As shown in the table, the Agency factor, regardless of the avatar type, improves with prior experience in virtual reality. This result indicates that the greater the prior experience with virtual reality technology, the stronger the sense of control over the avatar.

TABLE II
CORRELATION BETWEEN PRESENCE FACTORS AND VIRTUAL REALITY EXPERIENCE

<table>
<thead>
<tr>
<th></th>
<th>Property</th>
<th>Agency</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>0.252</td>
<td>0.469</td>
<td>0.327</td>
</tr>
<tr>
<td>N</td>
<td>0.191</td>
<td>0.329</td>
<td>0.192</td>
</tr>
<tr>
<td>O</td>
<td>-0.076</td>
<td>0.411</td>
<td>-0.062</td>
</tr>
</tbody>
</table>

B. Privacy

Regarding the sensation of a lack of privacy, users did not express significant concern in the initial three tests, registering a value below 0.5 on a scale of 0 to 4. However, this concern heightened during the fourth test when another person employed their hyper-realistic avatar (1.5). To investigate the correlation between the sensation of a lack of privacy and the sense of presence, a Spearman correlation analysis was performed, and the outcomes are detailed in Table III.

The table reveals a correlation between the three factors and the sensation of a lack of privacy, with a pronounced emphasis on hyper-realistic avatars and the sense of change (r=0.575; p=0.001), and to a lesser extent, the sense of control (r=0.397; p=0.014). Thus, a heightened sensation of change and control over the avatar corresponds to an increased apprehension about a lack of privacy. Furthermore, the table illustrates that when users utilize another person’s avatar and become aware that others can access their avatar, all three presence factors intensify this perception of a lack of privacy, establishing a clear relationship.

TABLE III
CORRELATION BETWEEN PRESENCE FACTORS AND PRIVACY FACTOR

<table>
<thead>
<tr>
<th></th>
<th>Property</th>
<th>Agency</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyper-realistic</td>
<td>0.215</td>
<td>0.397</td>
<td>0.575</td>
</tr>
<tr>
<td>Non-realistic</td>
<td>0.192</td>
<td>0.068</td>
<td>0.020</td>
</tr>
<tr>
<td>Other person</td>
<td>0.327</td>
<td>0.034</td>
<td>0.323</td>
</tr>
</tbody>
</table>

C. Feasibility

As Fig. 7 shows, users perceived the use of technologies with avatars in both education and leisure as useful (value greater than 2.5). However, they are still not certain about their use in the workplace, although they do find it quite useful (2.3). On the other hand, the intention to use is favorable (value greater than 2.5).

D. Avatar preference

Fig. 8 shows a comparison of avatar preferences in different setting. In the workplace setting, users prefer to use avatars of themselves and with the highest image quality possible. This
preference increases with the seriousness of the meetings. In the educational context, when it comes to the teacher, the same preference as in the workplace setting is observed. That is, users prefer an avatar of themselves with good image quality. However, in this case, the preference between hyper-realistic (47.6%) and realistic (42.9%) avatars is closer. But as students, the preference changes, still preferring an avatar of themselves, but the need for realism is not as high. However, for leisure activities, preferences change, and they also vary depending on the type of leisure. For socializing with friends or attending events, users prefer an avatar with good image quality. But in a setting where they have to interact with both friends and strangers, such as in games, users prefer non-realistic avatars (52.4%).

Fig. 8. Comparison of avatar preferences in different settings.

### IV. Conclusions

The present study aimed to investigate the use of different types of avatars in the workplace and educational environments. It sought to understand users’ acceptance of using avatars and their preferences regarding the type of avatar they would prefer. To achieve this goal, a video conference scenario was simulated using different avatars. These avatars included a hyper-realistic avatar of the user, a non-realistic avatar of the user, and a hyper-realistic avatar of another person. Additionally, a final test was conducted with another person using the user’s avatar to compare the sense of presence.

The study revealed valuable information about the use of avatars in different settings, namely, work, education, and leisure. The results showed that users have a highly positive intention to use avatars in work and educational contexts. However, it should be noted that the study’s participants came from the university environment. This suggests that avatars may be a promising tool for online meetings and training sessions. Another limitation of the study is that it only considered the user’s experience with their own avatar. Interaction with avatars as an interlocutor or assistant, for example, was not explored.

The theory that the sense of presence is crucial in avatar use has been confirmed by the results. As the sense of presence increases, users perceive avatars as more useful and express a greater intention to use them. Therefore, to maximize the benefits provided by avatars, it is essential to focus on increasing the sense of ownership and control and minimizing the perception of body changes in the avatar. As the results of this study have shown, avatars that closely resemble the user enhance the sense of ownership and control while reducing the sense of change.

It is important to highlight the relationship between prior experience with virtual reality technology and the sense of control and perceived utility and intention to use. This indicates that the more experience a user has, the more willing they are to use this technology.

Regarding privacy concerns, the results show that, in general, users have few privacy concerns when using avatars. However, this fear increased when users were informed that someone else could use their avatar. In other words, although the initial results indicate low concern, this concern may increase over time due to the perception of the privacy of the information used to create avatars that could be sold or stolen [34]. Therefore, addressing this issue will be crucial in any technology that uses avatars, mainly in contexts where confidentiality is crucial as in the work context. For example, by maintaining transparency in data use and security [35]–[37].

Based on the importance of presence and the results related to avatar types, it is concluded that the best type of avatar to enhance presence is a realistic representation of the user with good image quality. This aligns with user preferences in work and education settings, where users prefer avatars that allow them to be presented in the best possible quality.

In conclusion, avatars offer significant potential to enhance interactions and meetings using virtual reality, particularly in professional and educational settings. By focusing on increasing presence and addressing user concerns about appearance and privacy, avatars can become a valuable tool for improving virtual interactions and user experiences.

### REFERENCES
