Journal of Information Systems and Informatics
Vol. 6, No. 3, September 2024 e-ISSN: 2656-4882 p-ISSN: 2656-5935
DOI: 10.51519/journalisi.v6i3.842 Published By DRPM-UBD

Creating Realistic Human Avatars for Social Virtual Environments Using Photographic Inputs

Raymond Leonardo Chandra [1], Koen Castermans[2], Djamel Berkaoui[3], Patrick Querl [4], Nacken Heribert [5]

[1,2,3,4,5]Academic and Research Department Engineering Hydrology, RWTH Aachen University, Aachen, Germany
Email: [1]chandra@lfi.rwth-aachen.de, [2]castermans@lfi-rwth-aachen.de, [3]berkaoui@lfi-rwth-aachen.de, [4]querl@lfi-rwth-aachen.de, [5]nacken@lfi-rwth-aachen.de

Abstract

This paper presents the development and evaluation of realistic virtual reality avatars created with a Blender add-on called Facebuilder. In this process, a person's head is photographed from different angles. These photographs are used in subsequent steps to generate a realistic avatar face. To investigate the user experience of interacting with these avatars, a study was conducted in VR using the MyScore application. The study involved 22 participants who met in a virtual environment to discuss a topic of their choice. Statistical analyses including descriptive statistics, Wilcoxon Signed-Rank Test, and Friedman Test showed significant differences supporting all three hypotheses: users preferred communicating with realistic avatars, were more focused and engaged when interacting with them. The results indicate a significant preference for realistic avatars in educational use cases, primarily due to the perceived seriousness of the interactions and the resulting higher level of participant engagement. The suitability of realistic versus non-realistic avatars was found to be use-case dependent. Participants suggested that realistic avatars would be more appropriate for educational scenarios and non-realistic avatars for entertainment.

Keywords: Virtual Reality, Education, Open-Source Software, Photogrammetry, Avatar

1. INTRODUCTION

Over the past few years, virtual reality (VR) has grown rapidly, and with the advancement of technology, the possibility of experiencing a fully immersive digital world is becoming more feasible [1]. One of the main factors contributing to immersion in VR is the avatar, which is a digital representation of the user in a virtual environment (VE). The level of realism of these avatars has a significant impact on the user's experience, as it can affect how they perceive themselves and how they interact with other users in the VE [2]. As a result, there has been a growing interest in avatar realism in VR. This interest stems from the belief that the more realistic an avatar appears, the more immersive the experience will be [3].

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Creating a realistic avatar for VR can be a challenging and complex process that requires a number of different techniques and technologies. One effective way to create a realistic VR avatar is to completely scan a person in real life and then use the data to create a digital representation of the person [4]. This process, which is widely known as Photogrammetry, can be accomplished using a variety of scanning technologies such as 3D scanning, laser scanning, and depth cameras such as the Microsoft Kinect [5]. Photogrammetry is not entirely new and has been discussed in several previous articles. While high-fidelity realistic avatars generated from photogrammetric 3D scanning methods have been shown to enhance virtual body ownership compared to abstract avatars [6], these methods often require sophisticated setups involving multiple cameras or specialized equipment [7], [8], [9], [10]. These approaches can be expensive and demand expertise to operate properly [5] which make them less accessible for everyday users. In addition, the resulting 3D avatars are often high in polygon count, which can unnecessarily consume computing power, especially in untethered VR headsets such as the Meta Quest 1 or 2 [10]. To overcome this problem, the generated 3D avatar needs to be optimized by reducing the polygon count before it is imported into the VR application. Such processes require manual editing of the 3D avatar which consumes time.

To avoid these problems of high cost and high polygon count, the Department of Engineering Hydrology (LFI) at RWTH Aachen University uses an alternative photogrammetry technique using an add-on for the 3D modeling software Blender, called Facebuilder [12]. By using the Facebuilder add-on, the photogrammetry process is simplified by enabling the creation of realistic avatars with just any smartphone camera. This significantly lowers the barrier to entry and makes the technology accessible to the public. Additionally, Facebuilder offers several distinct advantages compared to standard photogrammetry applications when it comes to creating lifelike digital human faces and heads, particularly in terms of integration and lower hardware demands [12].

Traditional photogrammetry typically involves capturing a large number of images from multiple angles, which requires a sophisticated setup capable of simultaneously operating more than 10 cameras [13]. Once these images are captured, they are processed together to create a digital human face and head, a step that demands substantial computational resources [13]. This complexity and resource requirement can be expensive in traditional photogrammetry [14]. In contrast, Facebuilder integrates directly with Blender which enables immediate changes and precise control over the details of the face. Facebuilder only requires a set of photographs from a person to create a realistic representation of the person's face and head. These photos can simply be taken with a smartphone camera. Facebuilder furthermore offers preset settings for your model's polygon count and thus makes it easier to develop within the limitations of your targeted

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

hardware. These unique benefits make Facebuilder a more efficient and user-friendly tool for artists who need to create precise 3D avatars using only a few photos and less demanding hardware.

The primary research gap addressed by this study lies in evaluating whether Facebuilder-generated avatars can offer comparable levels of realism and user engagement. Previous studies have not extensively explored this streamlined approach nor its implications on user experience within educational contexts. The aim of this paper is to evaluate how the participants rate the use of realistic avatars generated by Facebuilder compared to the non-realistic avatars. Three hypotheses were formulated which were: H1: Users will prefer to communicate with a realistic avatar over a non-realistic avatar. H2: Users will be more focused and engaged when interacting with realistic avatars than with non-realistic avatars. H3: An avatar generated by the Facebuilder add-on is sufficiently realistic to increase user's immersion

2. METHODS

As explained briefly before, Photogrammetry is the process of creating a 3D model of an object or person by taking multiple photographs of it from different angles [15]. These photographs are then processed using specialized software to create a 3D model. To create the avatar, multiple photos of the person are taken from different angles, and the software creates a 3D model of the person based on these photos. To make this technique accessible to everyday users, we decided to use smartphone cameras to capture the person's face from different angles. No additional devices like RGB-D cameras, Kinect or 3D scanners are needed to create the realistic 3D avatar.

2.1. Softwares

To process the photos of our subject and create the corresponding 3D avatar, we use FaceBuilder, a Blender add-on designed to digitally reconstruct human heads and faces from images. Blender is an open-source 3D modeling and animation software that can be extended with add-ons to enhance its basic capabilities and provide additional tools or options to the user [16]. We chose Blender to develop our 3D avatars because it is an extensible open source software and has a large community that supports and educates users about its capabilities. FaceBuilder is a Blender add-on developed by KeenTools for the visual effects industry. It supports the creation and animation of realistic individual faces using AI-based face recognition [12]. Using a base model and AI technology, FaceBuilder automatically generates 3D models and textures of a digital face that can be exported to game engines such as Unity for further use in VR development.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

To test the user's perception of the avatar, a VR software called MyScore was used. MyScore is an open source VR software developed by the LFI of RWTH Aachen University [17]. MyScore enables students and teachers to communicate via avatar-based teaching and learning scenarios in virtual reality.

2.2. Avatar Creation Process

Achieving a certain level of photographic realism in virtual reality avatars is important to give users a sense of anthropomorphic recognition while communicating in a virtual space [18]. Figure 1 summarizes the avatar creation process, which involves generating a realistic head of a person along with an avatar body. This leads to the development of a fully rigged avatar in Unity for user interaction.

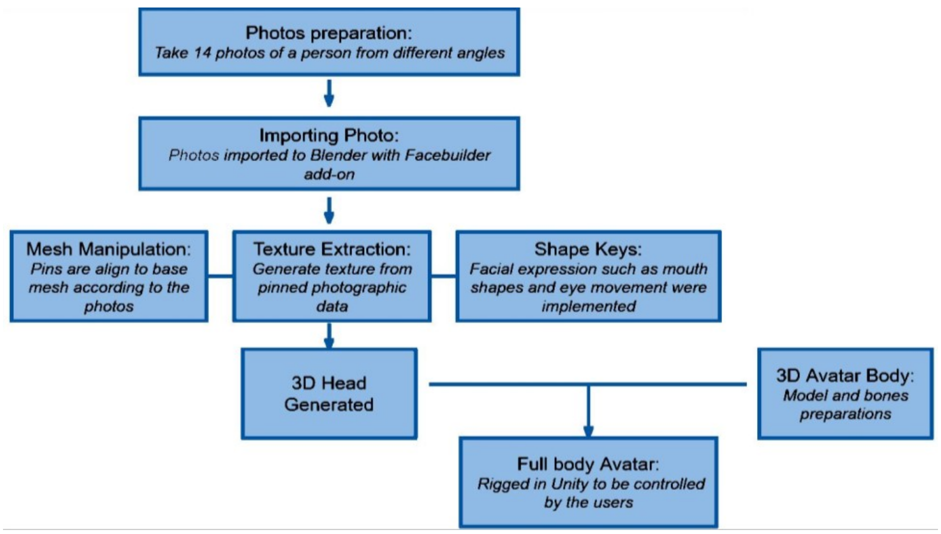


Figure 1. The flowchart of the realistic avatar creation process

To achieve this level of realism we collect fourteen photographs of our subject and process this data along with some manual refinement through the use of the Blender add-on Facebuilder. The photographs can be taken with everyday devices such as a smartphone's camera.

After considering all the important angles from which a face can be recognized, we decided that fourteen perspective images were needed. Figure 2 depicts the fourteen photographs needed. The angles include a downward, frontal, and upward view from the front and both sides of the head. In addition, a downward and upward photograph is taken from each three-quarter view. Finally, a single view of the back of the head is required. This number of views is important to get

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

enough information for the two main features of Facebuilder, namely mesh manipulation and texture extraction.

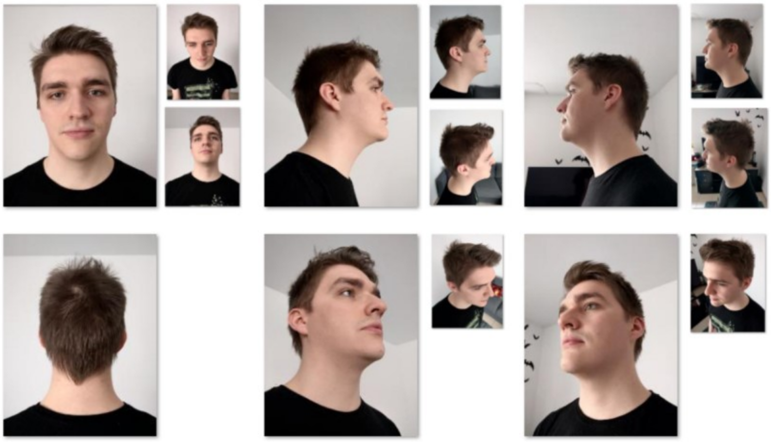


Figure 2. A set of photographs to be imported into Facebuilder

2.2.1. Mesh Manipulation

Initially, Facebuilder provides users with a basic geometric model, called a mesh, of a neutral, realistic human head (Figure 3A). This base mesh can be manipulated through a process called pinning. Pinning requires the user to align automatically or manually placed pins on the mesh according to the photographic information. By fine-tuning the location of these pins, the mesh is manipulated to resemble the real person in the photographs. Facebuilder allows the user to perform most of the pinning process in a matter of seconds using facial recognition. Afterwards, these automatic -and manually added pins, can still be manipulated by the user to ensure the quality of the end result.

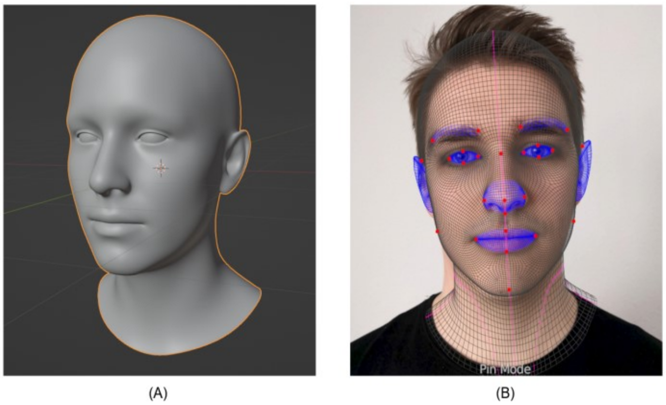


Figure 3. (A) Base Facebuilder Mesh. (B) Example of Pinning Process.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Figure 4 shows us an example of using the Facebuilder interface for processing our digital photos. The menu on the right shows all the photographs of our subject uploaded into Facebuilder. Additionally, it illustrates how each image corresponds to the current status of our mesh. Unlike the fully pinned example in Figure 3B, Figure 4 has not yet been aligned with our base mesh, as evidenced by the mismatch between the photo and the overlaying grid lines. Additionally, the Auto-Align button below the list of imported images utilizes facial recognition to speed up this process.

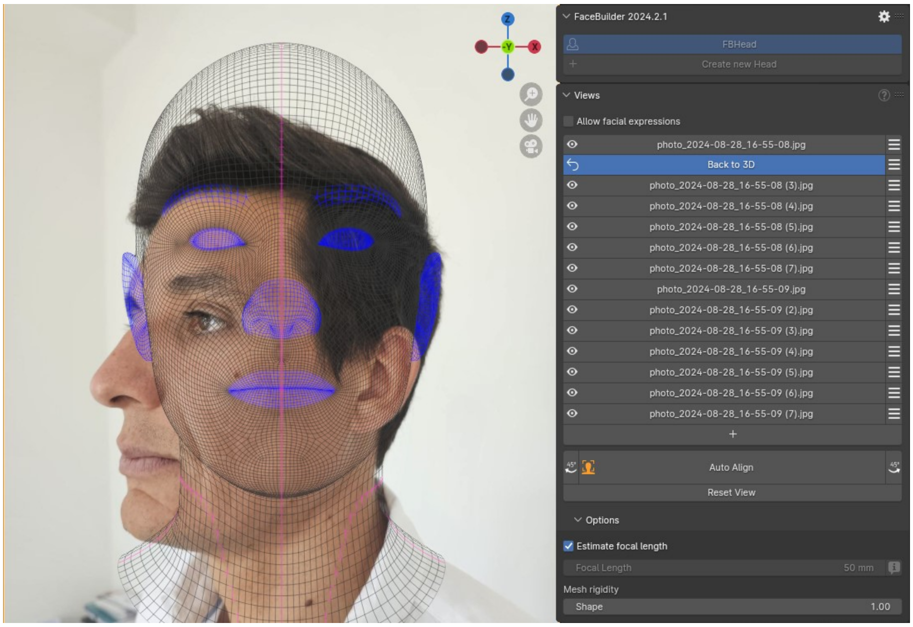


Figure 4. Processing of the digital imagery via the Facebuilder interface.

2.2.2. Texture Extraction

Once the pinning process was completed, the subsequent step, which was crucial in obtaining a satisfactory level of photographic realism, was texture extraction. During texture extraction, all the photographic information that was pinned onto the base mesh was used to generate a texture. To ensure a complete and seamless texture, it was imperative to have access to all fourteen photographs that captured the entire head of our subject. Otherwise, gaps may be present in the final texture. The end result consists of a composite texture that takes information from all the pinned captured images. This information is then merged to generate a new texture, like depicted in figure 5A, and projected correctly onto our mesh. If the photographic data is of high quality and the pinning process is well done, a satisfactory end result will be obtained as shown in figure 5B.

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

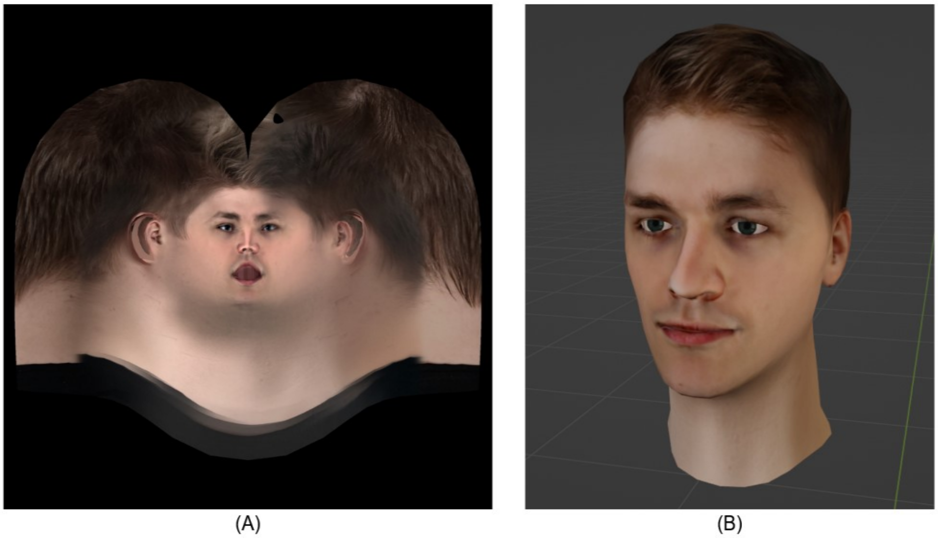


Figure 5. (A) Facebuilder Texture Result. (B) Finished Textured Model

2.2.3. Further Processing and Export to Unity

Before the avatar is imported into Unity for further use in VR development, the generated head is connected to an avatar rig and the facial expression data is implemented. Avatars with facial animation and body motion are generally perceived as more realistic [19]. Therefore, a certain type of animation data, specifically referred to as shape keys, which mimic facial expressions were implemented. These shape keys are based on the common mouth shapes and eye movements humans make while speaking. An algorithm was employed to activate these shape keys in real-time which translate the user's speech to avatar mouth movements by detecting sound emitted from the user. This gives the impression that the avatar was talking to the user.

Finally, the avatar is rigged with bones that track the user's movements in virtual reality during runtime. Rigging is a process in computer graphics that involves creating the digital skeleton that is needed to animate a 3D model or character. Each part of the virtual avatar's body corresponds to a bone. The bones use information from the VR controllers to determine the movement of each body part in virtual space. Figure 6A shows the bone structure of the human model. The generated realistic heads are then attached to the human model. Figure 6B shows the generated full body avatar with a realistic head.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

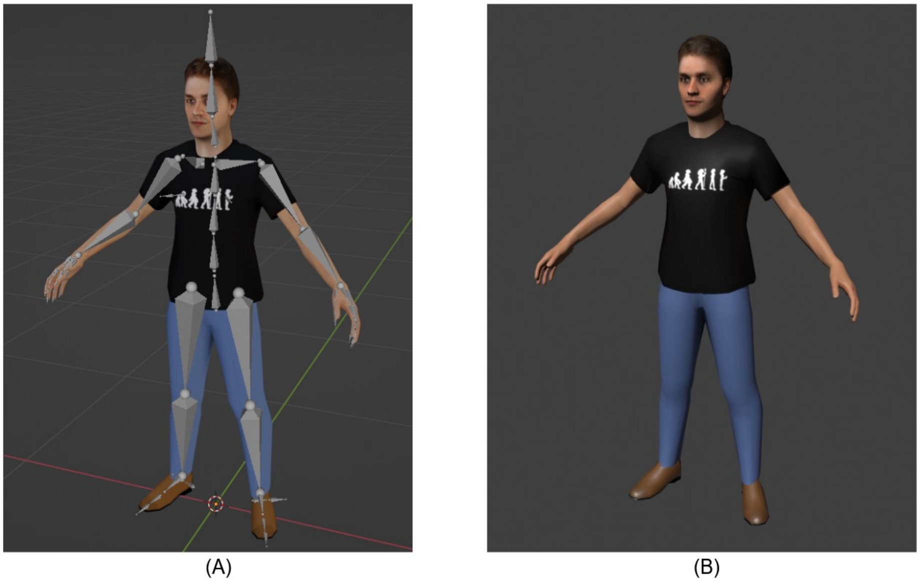


Figure 6. (A) The Underlying Rig Structure. (B) The Finished Avatar.

2.3. Experimental Design

The participants for this study joined voluntarily after being informed about the study's purpose and procedures, which ensure a diverse group of individuals. Each participant received detailed information about the study's objectives, what their participation would entail, and how their data would be used. This transparency ensured participants were aware of the commitment and felt comfortable taking part, which improves the study's reliability. Figure 7 outlines how the experiment was conducted.

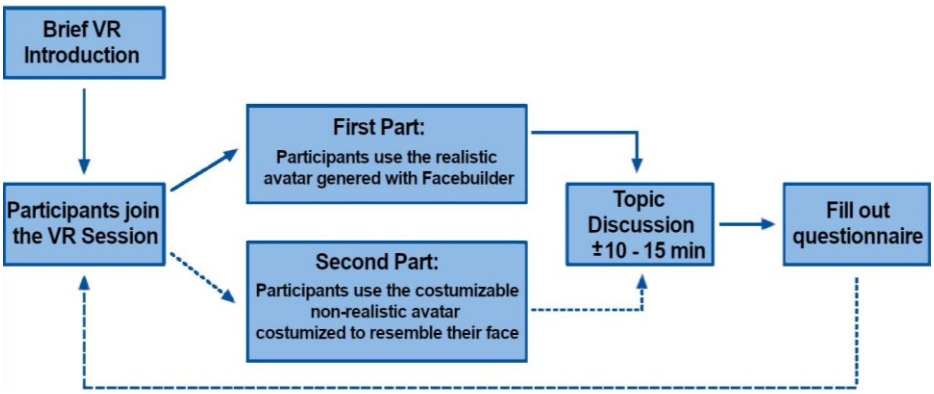
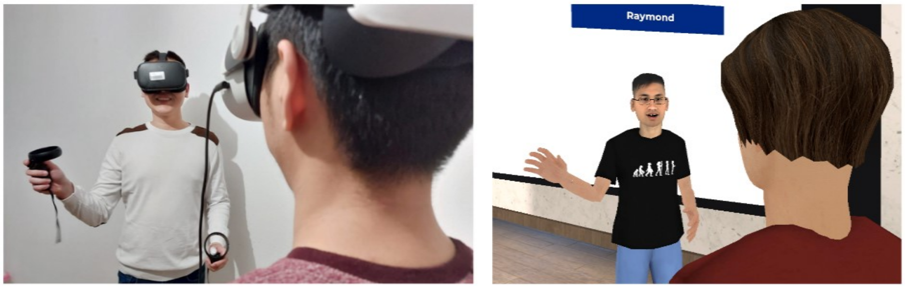


Figure 7. The flow chart of the experiment

Journal of Information Systems and Informatics
Vol. 6, No. 3, September 2024

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

The experiment began with a brief introduction to VR which covered essential VR controls, such as how to move and interact with objects within the VE. This introduction was required to accommodate participants with different levels of familiarity with VR technology. The experiments then continue with participants joining the virtual meeting session. To reduce distraction when judging avatar appearance, each session consisted of only two participants. Figure 8 shows the experiment taking place simultaneously in real life and in VR. In total two sessions were performed.



(A) (B)

Figure 8. Participants view in (A) real life and (B) VR

Participants were free to choose the topics they wanted to discuss. This freedom aims to create natural interactions and reduce bias that could arise from unfamiliar or uncomfortable topics. Each virtual meeting lasted 10-15 minutes. After the first part of the experiment, participants were asked to fill out a questionnaire rating how realistic the appearance of the avatar created with the Face Builder add-on was. Next, in the second part, participants were asked to choose a non-realistic humanoid avatar that they could customize themselves (see Figure 9).

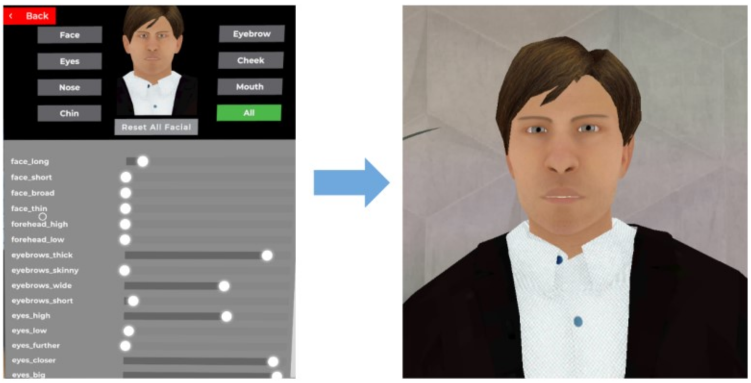


Figure 9. The participant's own non-realistic humanoid avatar

After re-entering the virtual meeting session and discussing freely for about the same amount of time as in the first part of the experiment, participants were again

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

given a series of questionnaires in which they rated the appearance of the realistic and non-realistic avatars on a Likert scale from 1 to 5. 1 being the lowest to 5 being the highest. They were also asked to rate the interaction with each other in a VE. Finally, they evaluated whether the use of a realistic avatar plays an important role in virtual immersion during a meeting.

2.4. Statistical Analysis

To analyze the three hypotheses mentioned in chapter 1, three statistical methods, specifically descriptive statistics, the Wilcoxon signed-rank test, and the Friedman test, were used to gain an understanding of the participants' preferences over realistic avatars as opposed to non-realistic ones. Furthermore, through the mentioned statistical methods, we can analyze whether the hypotheses H1 to H3 were statistically significant or not. Descriptive statistics provide insight to understand the basic features of the data through the average values of mean, median, mode [20]. This information provides an overview of what the general participants prefer and do not prefer.

The Wilcoxon Signed-Rank Test is a non-parametric test that is ideal for ordinal data, which does not assume a normal distribution [20]. This test was specifically designed for paired data, making it appropriate for comparing each participant's rating against a neutral point (3). In this experiment, the same participants provided ratings for both types of avatars in which a paired data was created for the needed analysis. The test helps determine if there were significant differences in participant preferences or engagement levels between the two types of avatars.

To further analyze the participants' responses, the Friedman Test, another non-parametric test, is particularly suitable to compare responses across multiple related questions within each hypothesis [20]. This test assesses whether there were significant differences in the ratings across different conditions, effectively handling the non-normal distribution and ordinal nature of the data. The test was also selected as it was appropriate to handle ordinal scales, which were used in this study's questionnaires [20]. By comparing responses across various conditions within each hypothesis, the Friedman Test helps identify significant differences in participant perceptions and experiences. This makes it suitable for evaluating comprehensive user feedback across multiple criteria.

The combination of these three statistical methods ensure comprehensive analysis of participants' ratings. Descriptive statistics provide a clear summary of the data, the Wilcoxon Signed-Rank Test identifies significant deviations from a neutral point, and the Friedman Test compares responses across multiple related questions, revealing significant differences in participants' perceptions of realistic avatars.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

3. RESULTS AND DISCUSSION

3.1. Participants

The study involved 22 participants, consisting of 17 men and 5 women. Their ages varied between 24 and 37 years old, with a mean age (μ) of 27 years old. Most of the participants (18 out of 22) were university students. There were 11 participants who knew about VR but they had not personally experienced it. A small group, 7 participants, had experienced VR, but rarely. For instance, they might have tried VR at a tech fair, used a colleague's VR headset, or tested it in a store, but they do not regularly engage with VR in their daily lives. There were 2 participants who used VR frequently and are likely to own a VR device and incorporate it into their regular routine. They might use VR for various purposes such as gaming or exploring virtual environments, making VR a significant part of their digital experience. The 2 remaining participants had never heard of VR before the experiment.

3.2. Statistical Result

Table 1 presents a detailed analysis of participant's preferences and engagement with realistic versus non-realistic avatars. The table includes data from 12 questions asked to the 22 participants. Each of the question were rated on a scale from 1 to 5, with 1 representing a preference for non-realistic avatars, 2 represents in-between nonrealistic and neutral preferences, 3 representing a neutral stance with no strong preferences towards both avatars, 4 represents the stance in-between a neutral and realistic avatar preference and 5 represents a preference for realistic avatars. This customized approach was chosen to directly address the unique aspects of the experimental setup and the specific hypotheses being tested, ensuring that the questions were closely aligned with the study's objectives. Therefore, the customized scale was employed to more accurately reflect the study's specific focus on the effectiveness of avatars created using the Facebuilder add-on. In addition, table 1 describes the three statistical methods used for analysis and whether the hypotheses were accepted or not. We used a free program called PSPP, which is similar to SPSS, for data analysis and calculations in this paper.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Table 1. Participant's feedback result and the result of statistical analysis

	No	Questions	Total participant who prefers:					Descriptive Statistic			Wilcoxon Signed-Rank Test		Friedman Test Analysis	
			A	**B**	**C**	**D**	**E**	Mean	Median	Modus	W-Statistic	p-value	F-Statistic	p-value
H1	1	Communicating is more serious with:	2	3	4	7	6	3.5	4	4	48	0.091	19.46	0.0002
	2	Communication is more enjoyable with:	1	1	6	6	8	4.1	4	5	17	0.002		
	3	I feel as if I am speaking with other real users just like in real life with:	2	1	7	6	6	3.6	4	4	45.5	0.073		
	4	Overall, you prefer communicating with:	2	1	6	5	8	3.7	4	5	26.5	0.026		
H2	5	Your level of focus improve with:	1	3	5	5	8	3.7	4	5	26.5	0.015	12.87	0.0059
	6	Level of engagement in the virtual environment increase with:	2	2	5	7	6	3.6	4	4	37	0.054		
	7	Personal space or distance is similar like in the real world with:	1	3	13	3	2	3.1	3	3	18.5	0.623		

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

	8	I prefer the immersion with other users during our interaction with:	0	2	5	7	8	4.0	4	5	10	0.001		
H3	9	The avatar's face looks realistic in:	1	1	4	7	9	4.0	4	5	18	0.002	12.87	0.0049
	10	The avatar's mouth movement matches speech in:	1	2	6	7	6	3.7	4	4	25	0.012		
	11	The avatar's expressions (like idle, smiling, raising eyebrow, etc.) feel realistic in:	2	4	7	5	4	3.2	3	3	45	0.379		
	12	I find the overall avatar is realistic in	0	2	7	6	7	3.8	4	3	9	0.003		

Note: The score of 1-5 which translates to A-E is only used in this table to reduce the confusion to the reader as to differentiate it with the participant counts. The weight remains the same (A equal to 1 and E equal to 5).
(A) prefers the non-realistic avatar; (B) is in between the non-realistic avatar and neutral; (C) indicates no strong preferences between the two avatars; (D) is between neutral and realistic avatar; and (E) prefer the realistic avatar
Acceptance value of = $p < 0.05$. Neutral Score= 3.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

The descriptive statistics analysis in Table 1 proves that all three hypotheses were accepted. The Q1 to Q4 which related to H1, show that most of the participants tended to choose realistic avatars over non-realistic avatars. For example, the mean ratings for these questions range from 3.5 to 4.1, with median values at 4 and modes frequently at 4 or 5. This suggests a tendency towards favoring realistic avatars for communication. Specifically, Q2 ("Communication is more enjoyable with:") has a mean of 4.1, indicating a strong preference for realistic avatars. H2 was proven to be true as participants in general also felt that the level of engagement or focus was higher when interacting with a realistic avatar. H2 was represented by the Q5 - Q8, which have a mean rating range from 3.6 to 4.0, with median and mode values indicating a preference for realistic avatars. For instance, Q5 ("Your level of focus improves with:") has a mean of 3.7, highlighting an enhanced focus with realistic avatars. Questions related to H3 (Q9 to Q12) also show that participants perceive Facebuilder-generated avatars as realistic. The mean ratings for these questions range from 3.7 to 4.0, with median values mostly at 4 and modes frequently at 4 or 5. This indicates that participants find the avatars' facial features, mouth movements, and overall realism satisfactory. For example, Q9 ("The avatar's face looks realistic in:") has a mean of 4.0, suggesting a strong perception of realism.

The Wilcoxon Signed-Rank Test compares participants' responses against a neutral point (3) and an acceptance value of $p < 0.05$ for each hypothesis. H1, significant differences were found in Q2 ($p = 0.002$) and Q4 ($p = 0.026$), indicating that participants find communication more enjoyable and overall preferable with realistic avatars. H2, significant differences were noted in Q5 ($p = 0.015$) and Q8 ($p = 0.012$), suggesting that participants' focus and immersion improve with realistic avatars. H3 was supported by significant differences in Q9 ($p = 0.002$), Q10 ($p = 0.012$), and Q12 ($p = 0.003$), indicating that participants perceive Facebuilder-generated avatars as realistic and capable of enhancing immersion. Overall the Wilcoxon Signed-Rank Test suggested that the three hypotheses were accepted.

The Friedman Test examines paired questionnaire data for each hypothesis, comparing it against a neutral point (score of 3) with a significance level of $p < 0.05$. The results indicate significant differences in participants' responses across questions within each hypothesis. The result shows that all the three hypotheses were also accepted. The participants exhibit significant preferences for communication with realistic avatars over non-realistic ones, as indicated in F-Statistic = 19.46, p-value = 0.0002. Moreover, participants' engagement and focus levels were reported to be significantly higher when interacting with realistic avatars compared to non-realistic ones, as shown by the value of F-Statistic = 12.87, p-value = 0.0059. Lastly, participants perceive avatars generated by the

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Facebuilder add-on as more realistic which enhance their sense of immersion, as indicated by the F-Statistic = 12.87, p-value = 0.0049).

There was additional info found from the statistical analysis that was worth mentioning as well. Most of the participants felt no significant difference in their perception of personal space when interacting with either type of avatar (as indicated by their responses to question 7). This evidence can be seen by the mean of 3.1 and the W-statistic of 18.5 and p-value of 0.62. However, when it came to the avatar's mouth movement (question 10), participants favored the realistic avatar (indicated by the mean of 3.7 and the W-statistic of 25 and p-value of 0.012). Interestingly, the same algorithm used to trigger the shape keys was implemented in both the realistic and non-realistic avatars. However, users still perceive the mouth movements of the realistic avatar to be more accurate than those of the non-realistic avatar. One possible interpretation of this data is that users, influenced by the perceived realism of the avatar's face, may be led to believe that the movements of the avatar's mouth corresponded more accurately to real-life speech. The avatar's expressions (question 11) for both non-realistic and realistic avatars employed the same algorithm, resulting in no significant difference in user perception (mean 3.2, W-statistic: 45, p-value: 0.37). This can be explained by the fact that no eye or mouth tracking was used to reliably translate this input from a real participant's expression to that of their avatar.

Based on the data found in table 1 we can conclude that all the three hypotheses can be accepted. The descriptive statistics indicate a general preference for realistic avatars, with mean ratings typically around 3.5 to 4. In addition, the Wilcoxon Signed-Rank Test shows significant differences from the neutral point for several questions, supporting the preference for realistic avatars in terms of communication enjoyment and engagement. Moreover, the Friedman Test confirms significant differences in responses across questions within each hypothesis, further validating the preference for realistic avatars for communication, engagement, and perceived realism. Overall, the analysis supports the acceptance of all three hypotheses, showing that participants favor realistic avatars over non-realistic ones.

Another question asked participants if they preferred avatars to be realistic or not in VR use cases. Figure 7 shows that four participants preferred a realistic avatar, while three participants preferred a non-realistic avatar. The remaining fifteen participants felt that the choice between realistic and non-realistic avatars should depend on the VR use case. The qualitative feedback indicated that participants found realistic avatars more suitable for educational scenarios due to their ability to enhance seriousness and concentration, while non-realistic avatars were deemed more appropriate for entertainment.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

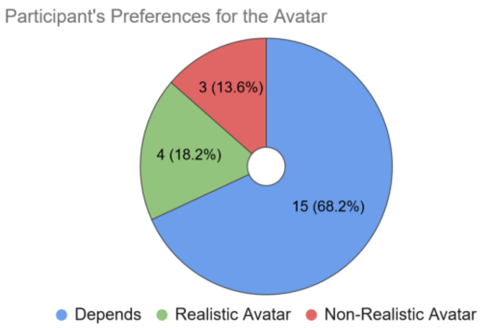


Figure 6. Participants preferences between the realistic and non-realistic avatar

This finding also correlates with the previous studies that have highlighted the significance of immersion in educational virtual reality, highlighting that realistic avatars enhance user engagement and foster a serious, focused atmosphere conducive to learning activities [1], [2]. On the contrary, the choice of avatar design can significantly influence the user's connection to the educational content and their overall learning experience. While realistic avatars enhance relatability and immersion, non-realistic avatars tend to emphasize fun and creativity, making them particularly appealing to younger audiences.

3.3. Discussion of Findings

The findings of this study indicate a significant preference for realistic avatars in educational VR environments, which aligns with previous literature [1]. This research underscores the importance of immersion in virtual reality as a pedagogical tool, emphasizing that realistic avatars can enhance user experience by creating a more engaging and serious atmosphere. This study's results corroborate these assertions, demonstrating that participants felt more focused and engaged when interacting with realistic avatars, thereby further supporting hypothesis H2. Furthermore, the discussion of the limitations and complexities associated with traditional photogrammetry techniques, is directly addressed by the use of the Facebuilder add-on in this study [9], [21]. By simplifying the photogrammetry process and reducing the need for specialized equipment, Facebuilder makes the creation of realistic avatars more accessible to the public, as discussed in the introduction. This approach aligns with the previous research which emphasized the need for more efficient and user-friendly photogrammetry methods [13].

Additionally, the preference for realistic avatars in communication and their perceived impact on immersion, as shown by the Wilcoxon Signed-Rank and

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

Friedman tests, aligns with previous findings that visual similarity between avatars and their users significantly enhances body ownership and perceived realism [2], [3]. This, in turn, leads to more effective and engaging interactions in virtual environments. The statistical significance found in this study further validates these claims, indicating that realistic avatars created with Facebuilder can indeed provide a more immersive VR experience.

In conclusion, the integration of Facebuilder for creating realistic avatars offers a practical solution to the challenges posed by traditional photogrammetry methods, while also enhancing user engagement and immersion in VR environments. These findings not only support the hypotheses tested but also align with the broader literature on the importance of avatar realism in virtual reality, suggesting significant potential for future applications in educational and professional settings.

4. CONCLUSIONS

The work presented here documents our creation of realistic virtual reality avatars. This process involves taking fourteen photographs of the subject from different angles and using those photographs to construct a realistic avatar. using the Facebuilder add-on within Blender. Facebuilder provides a base mesh that can be automatically, and manually, aligned to the photographic information via a process called pinning to generate a realistic face. Texture extraction is then used to achieve realism by projecting the photographic data onto the model. The generated face is connected to an avatar rig, enabling body movement and facial expressions using shape keys. Finally, the full-body avatar uses VR controllers to track the user's movements in virtual reality.

A study was conducted to evaluate the quality, and use, of the full-body avatars. Based on the existing literature and previous observations, three hypotheses were formulated regarding the significant role of realistically generated avatars. In the study, participants took part in VR-based face-to-face meetings where they first interacted with a realistic avatar, created using the Facebuilder add-on, followed by an interaction with a non-realistic avatar. Afterwards, they filled out a questionnaire about their preferences regarding the avatars. Three statistical analyses, including descriptive statistics, the Wilcoxon Signed-Rank, and the Friedman test, were used to analyze the data, indicating a clear preference for realistic avatars.

The findings have several practical implications for virtual reality applications. Realistic avatars are shown to improve communication and engagement levels in educational settings. Participants reported higher focus when interacting with realistic avatars, which shows the potential benefits for professional training and

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

remote collaboration as well. Additionally, using tools like Facebuilder is proven to make the creation of high-quality avatars more accessible without requiring expensive equipment or extensive expertise.
Outside educational contexts, these findings suggest that realistic avatars could enhance user experiences in diverse VR applications, for example telemedicine consultations where accurate representation can build trust between patients and doctors; corporate training programs where immersion can lead to better skill acquisition; social VR platforms where lifelike interactions could improve engagement among users; and entertainment industries where both types might serve different purposes depending on context (e.g., realism for narrative-driven experiences vs stylization for creative freedom).

For future work, the research in the field of virtual avatars could concentrate on improving the precision and authenticity of facial expressions in avatars. This could include incorporating facial feature tracking technologies to produce more natural animations which allows for exploration of how realistic avatars impact interactions across various fields such as healthcare, engineering, and social networking. Conducting long-term studies could also assess how sustained use of realistic versus non-realistic avatars impacts user experience over time.

5. ACKNOWLEDGEMENTS

The support of the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung - BMBF) & German Academic Exchange Service (Deutscher Akademischer Austauschdienst - DAAD) through the funding of the MyScore is gratefully acknowledged.

REFERENCES

[1] D. Hamilton, J. McKechnie, E. Edgerton, and C. Wilson, "Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design," *J. Comput. Educ.*, vol. 8, no. 1, pp. 1–32, Mar. 2021, doi: 10.1007/s40692-020-00169-2.

[2] D. Jo *et al.*, "The impact of avatar-owner visual similarity on body ownership in immersive virtual reality," in *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, Gothenburg Sweden: ACM, Nov. 2017, pp. 1–2. doi: 10.1145/3139131.3141214.

[3] E. Molina, A. R. Jerez, and N. P. Gomez, "Avatars rendering and its effect on perceived realism in Virtual Reality," in *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, Utrecht, Netherlands: IEEE, Dec. 2020, pp. 222–225. doi: 10.1109/AIVR50618.2020.00046.

[4] M. C. Fysh, I. V. Trifonova, J. Allen, C. McCall, A. M. Burton, and M.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Bindemann, "Avatars with faces of real people: A construction method for scientific experiments in virtual reality," *Behav. Res. Methods*, vol. 54, no. 3, pp. 1461–1475, Jun. 2022, doi: 10.3758/s13428-021-01676-5.

[5] D. Hepperle, C. F. Purps, J. Deuchler, and M. Wölfel, "Aspects of visual avatar appearance: self-representation, display type, and uncanny valley," *Vis. Comput.*, vol. 38, no. 4, pp. 1227–1244, Apr. 2022, doi: 10.1007/s00371-021-02151-0.

[6] 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 *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, Gothenburg Sweden: ACM, Nov. 2017, pp. 1–10. doi: 10.1145/3139131.3139156.

[7] Z. Li *et al.*, "3D Human Avatar Digitization from a Single Image," in *Proceedings of the 17th International Conference on Virtual-Reality Continuum and its Applications in Industry*, Brisbane QLD Australia: ACM, Nov. 2019, pp. 1–8. doi: 10.1145/3359997.3365707.

[8] A. Mao, H. Zhang, Y. Liu, Y. Zheng, G. Li, and G. Han, "Easy and Fast Reconstruction of a 3D Avatar with an RGB-D Sensor," *Sensors*, vol. 17, no. 5, p. 1113, May 2017, doi: 10.3390/s17051113.

[9] F. Bogo, M. J. Black, M. Loper, and J. Romero, "Detailed Full-Body Reconstructions of Moving People from Monocular RGB-D Sequences," in *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile: IEEE, Dec. 2015, pp. 2300–2308. doi: 10.1109/ICCV.2015.265.

[10] C. Pirch, O. Kausch, A. Klepser, and S. Morlock, "One Avatar, Two Level of Detail, One Result? - Analyzing the Effect of Low and High Detailed Avatars on Fitting Simulations," in *Proceedings of 3DBODY.TECH 2021 - 12th International Conference and Exhibition on 3D Body Scanning and Processing Technologies, Lugano, Switzerland, 19-20 October 2021*, Lugano, Switzerland: Hometrica Consulting - Dr. Nicola D'Apuzzo, Oct. 2021. doi: 10.15221/21.46.

[11] A. Bönsch, A. Feng, P. Patel, and A. Shapiro, "Volumetric Video Capture using Unsynchronized, Low-cost Cameras:," in *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, Prague, Czech Republic: SCITEPRESS - Science and Technology Publications, 2019, pp. 255–261. doi: 10.5220/0007373202550261.

[12] n. d KeenTools, "FaceBuilder for Blender." Accessed: Jan. 01, 2024. [Online]. Available: https://keentools.io/products/facebuilder-for-blender

[13] T. Boe and C. Carter, "Human Photogrammetry: Foundational Techniques for Creative Practitioners," *Int. J. Comput. Graph. Animat.*, vol. 10, no. 1, pp. 1–20, Jan. 2020, doi: 10.5121/ijcga.2020.10101.

[14] O. C. Quispe-Enriquez, J. J. Valero-Lanzuela, and J. L. Lerma, "Craniofacial 3D Morphometric Analysis with Smartphone-Based Photogrammetry,"

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Sensors, vol. 24, no. 1, p. 230, Dec. 2023, doi: 10.3390/s24010230.
[15] S. Nebel, M. Beege, S. Schneider, and G. D. Rey, "A Review of Photogrammetry and Photorealistic 3D Models in Education from a Psychological Perspective," *Front. Educ.*, vol. 5, p. 144, Aug. 2020, doi: 10.3389/feduc.2020.00144.
[16] T. Roosendaal, *The Essential Blender: Guide to 3D Creation with the Open-Source Suite Blender*, Blender Foundation, 2007.
[17] D. Berkaoui, R. Chandra, and K. Castermans, "MyScore – Avatar-Based Teaching and Learning," 2022, doi: 10.18420/VRAR2022_7790.
[18] K. L. Nowak, J. Fox, and The Ohio State University, "Avatars and computer-mediated communication: A review of the definitions, uses, and effects of digital representations on communication," *Rev. Commun. Res.*, vol. 6, pp. 30–53, 2018, doi: 10.12840/issn.2255-4165.2018.06.01.015.
[19] C. M. Grewe, T. Liu, C. Kahl, A. Hildebrandt, and S. Zachow, "Statistical Learning of Facial Expressions Improves Realism of Animated Avatar Faces," *Front. Virtual Real.*, vol. 2, p. 619811, Apr. 2021, doi: 10.3389/frvir.2021.619811.
[20] R. H. Riffenburgh, "Tests on Ranked Data," in *Statistics in Medicine*, Elsevier, 2012, pp. 221–248. doi: 10.1016/B978-0-12-384864-2.00011-1.
[21] Jing Tong, Jin Zhou, Ligang Liu, Zhigeng Pan, and Hao Yan, "Scanning 3D Full Human Bodies Using Kinects," *IEEE Trans. Vis. Comput. Graph.*, vol. 18, no. 4, pp. 643–650, Apr. 2012, doi: 10.1109/TVCG.2012.56.