

# VALID: A perceptually validated Virtual Avatar Library for Inclusion and Diversity

Tiffany D. Do  
tiffanydo@knights.ucf.edu  
University of Central Florida  
Orlando, Florida, USA

Mar Gonzalez-Franco  
margon@google.com  
Google Labs  
Seattle, Washington, USA

Steve Zelenty  
steve.zelenty@knights.ucf.edu  
University of Central Florida  
Orlando, Florida, USA

Ryan P. McMahan  
rpm@ucf.edu  
University of Central Florida  
Orlando, Florida, USA

## ABSTRACT

As consumer adoption of immersive technologies grows, virtual avatars will play a prominent role in the future of social computing. However, as people begin to interact more frequently through virtual avatars, it is important to ensure that the research community has validated tools to evaluate the effects and consequences of such technologies. We present the first iteration of a new, freely available 3D avatar library called the *Virtual Avatar Library for Inclusion and Diversity (VALID)*, which includes 210 fully rigged avatars with a focus on advancing racial diversity and inclusion. We present a detailed process for creating, iterating, and validating avatars of diversity. Through a large online study ( $n = 132$ ) with participants from 33 countries, we provide statistically validated labels for each avatar’s perceived race and gender. Through our validation study, we also advance knowledge pertaining to the perception of an avatar’s race. In particular, we found that avatars of some races were more accurately identified by participants of the same race.

## 1 INTRODUCTION

The study of human behavior in virtual worlds is becoming increasingly important as more and more people spend time in these digital spaces. In particular, understanding how behaviors in virtual environments compare to behaviors in real life can provide insight into the ways in which humans interact with technology and each other. Furthermore, it can also help inform the design of virtual environments, making them more realistic and socially acceptable.

One area of significant interest is virtual avatars, which are 3D representations of virtual humans often used in virtual worlds and simulations [59]. Virtual avatars have a prominent role in social computing and immersive environments [18, 28, 74]. For example, previous studies have utilized virtual avatars to express emotions [11], recreate social psychology phenomena [17, 25], and influence feelings of presence or embodiment [52, 55]. Moreover, virtual avatars are expected to play evermore important roles in the future. For instance, virtual avatars are at the center of every social virtual reality (VR) and augmented reality (AR) interaction, as they are key to representing remote participants [60] and facilitating collaboration [69].

Open virtual avatar libraries facilitate research by providing freely available resources to the community. For instance, Microsoft’s Rocketbox avatar library [24] has been extensively used in multiple studies (e.g., [13, 29, 32]), with over 400 stars and 100 forks on

GitHub since its release in 2020. However, the representation of races in currently available avatar libraries are limited and have not been validated through user perception studies. As examples, over 40% of the Rocketbox avatars are white males, and MakeHuman<sup>1</sup>, another popular resource for 3D avatars, only includes options for African, Asian, and Caucasian avatars. Due to such limitations, many researchers are unable to properly investigate the effects of diversity for virtual avatars. Furthermore, existing libraries not only have limited diversity but are also biased in their lack of inclusivity with regard to professions for minority avatars. For instance, Rocketbox lacks Asian avatars dressed for medical scenarios.

The limited representation of currently available resources may have an inherently detrimental impact by introducing biases into research investigations from the outset, such as restricting the development of scenarios with Asian medical professionals. Furthermore, as immersive technologies grow in popularity worldwide, Taylor et al. [70] have called for improved racial/ethnic representations in VR to mitigate potential negative effects of racial bias. This can be especially important since virtual avatar race<sup>2</sup> can have a strong influence on human behavior. For example, research has found that embodying a virtual avatar of a different race can affect implicit racial bias, as demonstrated by several prior studies [26, 42, 57, 63]. As such, it is critical to provide more inclusive avatar resources.

The goal of this work was dual purpose. First, we sought to develop an open library of validated virtual avatars, which we refer to as the *Virtual Avatar Library for Inclusion and Diversity (VALID)*. This library has been designed to inclusively represent a range of races across various professions. Secondly, we wanted to better understand how humans perceive avatars, particular with regard to race. In this first iteration of VALID, we followed the recommendations of the 2015 National Content Test Race and Ethnicity Analysis Report from the U.S. Census Bureau [45] to ensure that a wide range of people was represented in our library. For example, this report found that members of Middle Eastern and North African (MENA) communities use the MENA category when it is available, but have trouble identifying their race when it is not available. Similarly, it also found that Hispanics can better identify their ethnicity by combining the conventional race and Hispanic origin questions into one. Therefore, to overcome those traditional problems on racial

<sup>1</sup><http://www.makehumancommunity.org/>

<sup>2</sup>We use the term race to refer to a “visually distinct social group with a common ethnicity” [61].

and ethnic categorization, VALID includes seven races as recommended by the U.S. Census Bureau report [45] (which differs from the 2020 U.S. Census): American Indian or Native Alaskan (AIAN)<sup>3</sup>, Asian, Black or African American (Black), Hispanic, Latino, or Spanish (Hispanic), Middle Eastern or North African (MENA), Native Hawaiian or Pacific Islander (NHPI), and White.

VALID includes 210 fully rigged virtual avatars designed to advance diversity and inclusion. We iteratively created 42 base avatars (7 target races  $\times$  2 genders  $\times$  3 individuals) using a process that combined data-driven average facial features with extensive collaboration with representative stakeholders from each racial group. To address the longstanding issue of the lack of diversity in virtual designers and to empower diverse voices [9], we adopted a participatory design method. This approach involved actively involving individuals ( $n = 22$ ) from diverse backgrounds, particularly different racial and ethnic identities, in the design process. By including these individuals as active participants, we aimed to ensure that their perspectives, experiences, and needs were considered and incorporated into the design of the avatars.

Once the avatars were created, we sought to evaluate their perception on a global scale. We then conducted a large online study ( $n = 132$ ) with participants from 33 countries, self-identifying as one of the seven represented races, to determine whether the race and gender of each avatar are recognizable, and therefore validated. We found that all Asian, Black, and White avatars were universally identified as their modeled race by all participants, while our AIAN, Hispanic, and MENA avatars were typically only identified by participants of the same race, indicating that participant race can bias perceptions of a virtual avatar’s race. We have since modeled the 42 base avatars in five different outfits (casual, business, medical, military, and utility), yielding a total of 210 fully rigged avatars.

To foster diversity and inclusivity in virtual avatar research, we are making all of the avatars in our library freely available to the community as open source models. In addition to the avatars, we are also providing statistically validated labels for the race and gender of all 42 base avatars. Our models are available in FBX format, are compatible with previous libraries like Rocketbox [24], and can be easily integrated into most game engines such as Unity and Unreal. Additionally, the avatars come equipped with facial blend shapes to enable researchers and developers to easily create dynamic facial expressions and lip-sync animations. All avatars, labels, and metadata can be found at our GitHub repository: <https://github.com/xrtlab/Validated-Avatar-Library-for-Inclusion-and-Diversity---VALID>.

This paper makes three primary contributions:

- (1) We provide 210 openly available, fully rigged, and perceptually validated avatars for the research community, with a focus on advancing diversity and inclusion.
- (2) Our diversity-represented user study sheds new light on the ways in which people’s own racial identity can affect their perceptions of a virtual avatar’s race. In our repository, we also include the agreement rates of all avatars, disaggregated by every participant race, which offers valuable insights into how individuals from different racial backgrounds perceive our avatars.

- (3) We describe a comprehensive process for creating, iterating, and validating a library of diverse virtual avatars. Our approach involved close collaboration with stakeholders and a commitment to transparency and rigor. This could serve as a model for other researchers seeking to create more inclusive and representative virtual experiences.

## 2 RELATED WORK

In this section, we describe how virtual avatars are used within current research in order to highlight the need for diverse avatars. We conclude the section with a discussion on currently available resources used for virtual avatars and virtual agents.

### 2.1 Effect of Avatar Race

Virtual avatars are widely used in research simulations such as training, education, and social psychology. The race of a virtual avatar is a crucial factor that can affect the outcomes of these studies. For example, research has shown that underrepresented students often prefer virtual instructors who share their ethnicity [3, 48]. Similarly, studies have suggested that designing a virtual teacher of the same race as inner-city youth can have a positive influence on them [4], while a culturally relevant virtual instructor, such as an African-American instructor for African-American children, can improve academic achievement [21].

The design of virtual avatars is especially important for minority or marginalized participants. Kim and Lim [37] reported that minority students who feel unsupported in traditional classrooms develop more positive attitudes towards avatar-based learning. In addition, children with autism spectrum disorder treat virtual avatars as real social partners [1, 38]. Therefore, to better meet the needs of all individuals participating in such studies, it is important for researchers to have access to diverse avatars that participants can comfortably interact with. Diversity in virtual avatars is important not only for improving representation, but also for enhancing the effectiveness of simulations. Halan et al. [27] found that medical students who trained with virtual patients of a particular race demonstrated increased empathy towards real patients of that race. Similarly, Bickmore et al. [6] showed that interacting with a minority virtual avatar reduced racial biases in job hiring simulations.

These findings highlight the importance of diverse and inclusive virtual avatars in research simulations and emphasize the need for more comprehensive representation of different races. Access to a wide range of validated avatars through VALID will help to create more inclusive and representative simulations, and enable researchers to investigate the impact of avatar race or gender on participants’ experiences. This will help improve the inclusivity of simulations and contribute towards addressing issues of bias.

### 2.2 Implicit Racial Bias and Virtual Avatars

Avatars are becoming increasingly important in immersive applications, particularly in the realm of VR, where they are becoming ubiquitous [16]. Research suggests that the degree of similarity between a user’s real body and their virtual avatar can influence embodiment, presence, and cognition [34, 35, 46, 53]. Maister et al. [42] proposed that non-matching self-avatars can affect a user’s self-association and influence social cognition. Moreover, studies

<sup>3</sup>We use racial abbreviations as defined in the U.S. Census Bureau report.

have demonstrated that embodying a darker-skinned avatar in front of a virtual mirror can reduce implicit racial biases [26, 42, 57, 63], which are unconscious biases that can lead to discriminatory behavior [12]. For instance, Salmanowitz et al. [63] found that a VR participant’s implicit racial bias affects their willingness to convict a darker-skinned suspect based on inconclusive evidence. Similarly, Peck et al. [56] found that each participant’s implicit racial bias was related to their nuanced head and hand motions in a firearm simulation. These foundational studies provide compelling evidence that embodying an avatar of a different race can affect implicit biases and further emphasize the need for diverse avatar resources.

Our study examines how participants perceive the race of diverse virtual avatars. While some studies have explored how a virtual avatar’s race affects user interactions (e.g., [3, 19, 75]), little research has been conducted on how individuals *actively perceive* the race of virtual avatars. Setoh et al. [66] note that racial identification can predict implicit racial bias, making it crucial to understand how people perceive the race of virtual avatars to further investigate these effects.

### 2.3 Own-Race Bias

Own-race bias, also known as the "other-race effect," refers to the phenomenon in which individuals process the faces of their own race differently from those of other races [14, 20, 47, 62]. Studies have suggested that this bias can influence the way individuals categorize race. For example, MacLin and Malpass [41] found that Hispanic participants were more likely to categorize Hispanic faces as fitting their racial category than Black faces, and Blascovich et al. [7] observed that participants who strongly identify with their in-group are more accurate in identifying in-group members.

Although own-race bias has not yet been studied in the context of 3D virtual avatars, Saneyoshi et al. [64] recently discovered that it extends to the uncanny valley effect [49] for 2D computer-generated faces. Specifically, they found that Asian and European participants rated distorted faces of their own race as more unpleasant than those of other races. Building on this research, we extended the study of own-race bias to 3D virtual avatars and focused on race categorization rather than perceived pleasantness. Our study included avatars and participants from seven different races, providing insights into how a diverse user population may interact within equally diverse virtual worlds.

### 2.4 Virtual Avatar Resources

There are numerous resources for creating virtual avatars. Artists can use 3D modeling tools, such as Autodesk 3ds Max<sup>4</sup>, Autodesk Maya<sup>5</sup>, Blender<sup>6</sup>, or ZBrush<sup>7</sup> to manually model, texture, and rig virtual avatars. However, such work requires expertise in 3D modeling and character design, and is often a tedious process [24]. On the other hand, parametric models, including freely available tools like

MakeHuman<sup>8</sup> and Autodesk Character Generator<sup>9</sup>, as well as commercially available ones such as Daz3D<sup>10</sup>, Poser<sup>11</sup>, and Reallusion Character Creator<sup>12</sup>, enable users to generate virtual avatars from predefined parameters, thereby significantly expediting the avatar generation process. Nonetheless, using these tools still requires learning a new program and time to customize each model, despite the absence of the artistic expertise needed for manual tools.

Another alternative to traditional modeling is to use scanning technologies, which can capture 3D models of real people. For instance, Shapiro et al. [67] and Waltemate et al. [73] used 3D cameras and photogrammetry, respectively, to capture 3D models of their users. Singular Inversions FaceGen Modeller<sup>13</sup> has also been employed to generate 3D faces from user photos and then apply them to a general 3D avatar body [8, 22]. However, scanning approaches require the ability to physically scan the user, limiting their use for certain applications, particularly remote ones.

Most closely related to our goal of providing a free and open library of ready-to-use avatars is the Microsoft Rocketbox library [24] and its accompanying HeadBox [72] and MoveBox [23] toolkits. Rocketbox provides a free set of 111 fully rigged adult avatars of various races and outfits. However, it falls short in terms of representation by not including any avatars of AIAN or NHPI descent. Additionally, the library offers only a limited number of Asian, Hispanic, and MENA avatars, excluding minority representations for some professions (e.g., Rocketbox does not include any Asian medical avatars). Furthermore, none of the available avatar libraries have been validated by user perception studies to ensure their efficacy and inclusivity. Therefore, our VALID project aims to fill this gap by providing a free and validated library of diverse avatars.

## 3 AVATAR CREATION PROCEDURE

This section outlines our iterative process for developing the VALID library, which includes 42 base avatars. We began by using data-driven averaged facial features to create our initial models. We then conducted interviews with representative volunteers to iteratively refine and modify the avatars based on their feedback.

### 3.1 Initial Modeling

To ensure a broad diversity of people were represented in our library, we initially created 42 base avatars (7 target races  $\times$  2 genders  $\times$  3 individuals) modeled after the seven racial groups recommended by the 2015 National Content Test Race and Ethnicity Analysis Report [45]: AIAN, Asian, Black, Hispanic, MENA, NHPI, and White. We created 3 male and 3 female individuals for each race, resulting in a total of 6 individuals per race.

Preliminary models were based on averaged facial features of multiple photos selected from the 10k US Adult Faces Database [2] and stock photos from Google for races missing from the database (e.g., AIAN, MENA, and NHPI). These photos were used as input to a face-averaging algorithm [15], which extracted average facial features for each race and gender pair. Using these averages as

<sup>4</sup><https://www.autodesk.com/products/3ds-max/overview>

<sup>5</sup><https://www.autodesk.com/products/maya/overview>

<sup>6</sup><https://www.blender.org>

<sup>7</sup><https://www.maxon.net/en/zbrush>

<sup>8</sup><http://www.makehumancommunity.org/>

<sup>9</sup><https://charactergenerator.autodesk.com/>

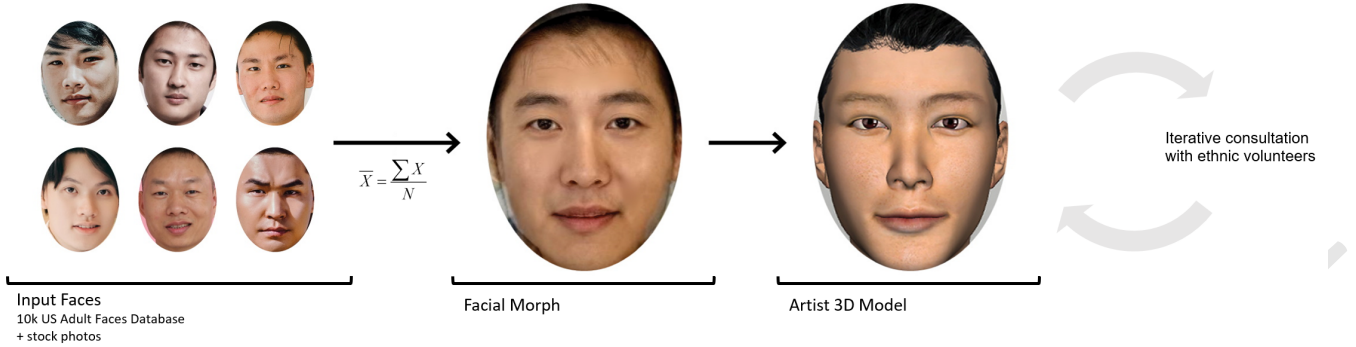
<sup>10</sup><https://www.daz3d.com>

<sup>11</sup><https://www.posersoftware.com>

<sup>12</sup><https://www.reallusion.com/character-creator/>

<sup>13</sup><https://facegen.com/modeller.htm>

## Pipeline



**Figure 1: An example of the creation of a 3D avatar using our methodology. 1) We select 4-7 faces from a database [2] or stock photos. 2) We calculate the average face using WebMorph [15]. 3) A 3D artist recreates the average face using modeling software. 4) The models are improved iteratively through recurrent consultation with representative volunteers.**

a reference, a 3D artist recreated the average faces for each race and gender pair using Autodesk Character Generator (due to its generous licensing and right to freely edit and distribute generated models<sup>14</sup>) and Blender to make modifications not supported by Autodesk Character Generator (see Figure 1).

### 3.2 Iterative Improvements through Representative Interviews

After the preliminary avatars were created based on the facial averages, we worked closely with 2 to 4 volunteers of each represented race (see Table 1) to adjust the avatars through a series of Zoom meetings. This process ensured that all avatars were respectful and reduced the likelihood of harmful or stereotypical representations. Volunteers self identified their race and were recruited from university cultural clubs (e.g., Asian Student Association, Latinx Student Association), community organizations (e.g., Pacific Islanders Center), and email lists. We iteratively asked these volunteers for feedback on all avatars representing their race, showing them the model from three perspectives (see Figure 2). Volunteers were specifically asked to identify accurate features and suggest changes to be made. Once the changes were completed based on the feedback, we presented the updated avatars to the volunteers. This process was repeated until they approved the appearance of the avatars. For example, volunteers requested changes to facial features, such as:

- “Many Native women I know have a softer face and jawline [than this avatar].” –ALAN Volunteer 3
- “The nose bridge is too high and looks too European. Asians mostly have low nose bridges.” –Asian Volunteer 2
- “Middle Eastern women usually have wider, almond-shaped eyes.” –MENA Volunteer 2
- “The nose [for this avatar] should be a little bit thicker, less pointy, and more round.” –NHPI Volunteer 1

<sup>14</sup><https://knowledge.autodesk.com/support/character-generator/learn-explore/caas/sfdcarticles/sfdcarticles/Character-Generator-Legal-Restrictions-and-Allowances-when-using-Character-Generator.html>

Additionally, we modified hairstyles according to feedback:

- “[These hairstyles] look straighter and more Eurocentric. So I would choose [these facial features] and then do a natural [hair] texture.” –Black Volunteer 1
- “Usually the men have curly hair or their hair is cut short on the sides with the top showing.” –NHPI Volunteer 1

Once the avatars were approved by their corresponding volunteer representatives, we conducted an online study to validate the race and gender of each avatar based on user perceptions.

**Table 1: Breakdown of our volunteer representatives by race, gender (male, female, or non-binary), and country.**

Race	Gender	Country
AIAN	2M, 1F	USA (Native American) (3)
Asian	2M, 2F	China (1), USA (Chinese-American) (1), USA (Vietnamese-American) (2)
Black	1M, 2F	USA (African-American) (3)
Hispanic	2M, 1F	Mexico (1), USA (Mexican-American) (2)
MENA	1M, 2F	Iran (2), Saudi Arabia (1)
NHPI	1M, 2F	Samoa (2), USA (Native Hawaiian) (1)
White	2M, 1F	USA (3)

## 4 AVATAR VALIDATION STUDY

We conducted an online, worldwide user study to determine whether the target race and gender of each avatar is recognizable and, therefore, validated. Participants were recruited from the online Prolific marketplace<sup>15</sup>, which is similar to Amazon Mechanical Turk. Prior research shows that Prolific has a pool of more diverse and honest participants [58] and has more transparency than Mechanical Turk [54]. Since diversity was a core theme of our research, we chose Prolific to ensure that our participants would be diverse.

<sup>15</sup><https://www.prolific.co>

Validation



**Figure 2: An example of how each avatar was presented to participants during our validation study.**

### 4.1 Procedure

The following procedure was reviewed and approved by our university Institutional Review Board (IRB). The study consisted of one online Qualtrics survey that lasted an average of 14 minutes. Each participant first completed a background survey that captured their self-identified demographics, including race, gender, and education. Afterwards, they were asked to familiarize themselves with the racial terms as defined by the U.S. Census Bureau research [45]. Participants were then asked to categorize the 42 avatars by their perceived race and gender. Participants were shown only one avatar at a time and the order was randomized.

For each of the avatars, participants were shown three perspectives: a 45° left headshot, a direct or 0° headshot, and a 45° right headshot (see Figure 2). Avatars were shown from the shoulders up and were dressed in a plain gray shirt. The images were rendered in Unity using the standard diffuse shader and renderer. The avatars were illuminated by a soft white (#FFFEF5) directional light with an intensity of 1.0, and light gray (#7F7F7F) was used for the background. Participants were asked to select all races that each avatar could represent: “American Indian or Alaskan Native”, “Asian”, “Black or African American”, “Hispanic, Latino, or Spanish”, “Middle Eastern or North African”, “Native Hawaiian or Pacific Islander”, “White”, or “Other”. “Other” included an optional textbox if a participant wanted to be specific. We allowed participants to select multiple categories according to the U.S. Census Bureau’s recommendations for surveying race [45]. For gender, participants were able to select “Male”, “Female”, or “Non-binary”. Participants were paid \$5.00 via Prolific for completing the study.

### 4.2 Participants

A total of 132 participants (65 male, 63 female, 4 non-binary) from 33 different countries were recruited to take part in the study. We aimed to ensure a diverse representation of perspectives by balancing participants by race and gender. Table 2 provides a breakdown of our participants by race, gender, and country. Despite multiple recruitment attempts, including targeted solicitations via Prolific, we had difficulty recruiting NHPI participants. It is important to note that we excluded volunteers who had previously assisted with modeling the avatars from participating in the validation study to avoid potentially overfitting their own biases.

### 4.3 Data Analysis and Labeling Approach

To validate the racial identification of our virtual avatars, we used Cochran’s Q test [68], which allowed us to analyze any significant

**Table 2: Breakdown of our validation study’s participants by race, gender (male, female, or non-binary), and country.**

Race	Gender	Country
AIAN	9M, 10F, 1NB	U.S. (14), Canada (3), Chile (1), Mexico (1), Cambodia (1)
Asian	10M, 10F	U.K. (5), Canada (3), South Africa (3), India (2), Indonesia (2), China (1), France (1), Germany (1), Italy (1), Malaysia (1)
Black	10M, 9F, 1NB	South Africa (16), Nigeria (2), Swaziland (1), U.K. (1)
Hispanic	10M, 9F, 1NB	Mexico (15), Chile (3), Portugal (2)
MENA	10M, 10F	Israel (9), Lebanon (3), Jordan (2), Bahrain (1), Egypt (1), Iran (1), Iraq (1), Saudi Arabia (1), Syria (1)
NHPI	7M, 5F	New Zealand (Maori) (8), Samoa (2), Fiji (1), U.S. (1)
White	9M, 10F, 1NB	Portugal (5), Italy (5), Poland (3), Mexico (2), Belgium (1), Greece (1), Ireland (1), U.K. (1), U.S. (1)

differences among the selected race categories. This approach was necessary since our survey format allowed participants to select more than one race category for each avatar, following the U.S. Census Bureau’s research recommendations [45]. Since the Chi-squared goodness of fit test requires mutually exclusive categories, we were unable to use it in our analysis. Furthermore, since our data was dichotomous, a repeated-measures analysis of variance (ANOVA) was not appropriate. Therefore, Cochran’s Q test was the most appropriate statistical analysis method for our survey data.

We used a rigorous statistical approach to assign race and gender labels to each avatar. First, we conducted the Cochran’s Q test across all participants ( $n = 132$ ) at a 95% confidence level to identify significant differences in the participants’ responses. If the test indicated significant differences, we performed pairwise comparisons between each race using Dunn’s test to determine which races were significantly different.

For each avatar, we assigned a race label if the race was selected by the majority of participants (i.e., over 50% of participants selected it) and if the race was selected significantly more than other race choices and not significantly less than any other race. This approach resulted in a single race label for most avatars, but some avatars were assigned multiple race labels due to multiple races being selected significantly more than all other races. If no race was selected significantly more than the majority, then we categorized the avatar as “Ambiguous”. We followed a similar procedure for assigning gender labels.

To account for the possibility that the race of the participant might influence their perception of virtual race, we also assigned labels based on same-race participants. This involved using the same procedure for assigning labels as described above, except based only on the selections of participants who identified as the same race as the avatar. This also allows future researchers to have the flexibility to use the labels from all study participants for studies focused on individuals from diverse racial backgrounds or to use the labels from participants of the same race for studies targeting specific racial groups.

## 5 RESULTS

### 5.1 Validated Avatar Labels

Table 3 summarizes our results and labels for all 42 base avatars across all participants and for same-race participants.

**5.1.1 Race and Gender Labels.** Asian, Black, and White avatars were correctly identified as their intended race across all participants, while most of the remaining avatars were accurately identified by same-race participants (see Table 3 for all and same-race agreement rates). Therefore, we observed some differences in identification rates based on the race of the participants, highlighting the potential impact of own-race bias on the perception of virtual avatars. Notably, there were no significant differences in gender identification rates based on participant race, indicating that all avatars were correctly perceived as their intended gender by all participants, regardless of their racial background.

**5.1.2 Naming Convention.** If an avatar was identified as its intended race by corresponding same-race participants, we named it after that race. For instance, the avatar `Hispanic_M_2` was labeled as White by all participants. However, our Hispanic participants perceived it as solely Hispanic. Hence, we left the original name. However, if an avatar was labeled as “Ambiguous” or as a different race by same-race participants, we added an X at the beginning of its name to indicate that it was not validated. Avatars were also labeled by their identified gender (“M” or “F”).

### 5.2 Other-Race vs. Same-Race Perception

To further examine how participant race affected perception of virtual avatar race, we additionally analyzed the data by separating same-race and other-race agreement rates. In effect, we separated the selections of the participants who were the same race as the avatar modeled and those who were not.

**5.2.1 Difference in Agreement Rates.** Figure 3 displays the difference in agreement rates between same-race and other-race. Figure 3 shows that several avatars were strongly identified by both other-race and same-race participants. In particular, all Asian, Black, and White avatars were perceived as their intended race with high agreement rates by both same-race and other-race participants (over 90% agreement for all but one). However, some avatars were only identified by participants of the same race as the avatar. For example, our analysis of the agreement rates for different racial groups revealed interesting trends. For instance, non-Hispanic participants had an average agreement rate of 54.5% for Hispanic avatars, while Hispanic participants had a much higher average agreement of 75.0%. Similar patterns were observed for AIAN (57.8% other-race, 75.0% same-race) and MENA (40.4% other-race, 68.0% same-race) avatars.

**5.2.2 Perceived Race Clusters.** To gain deeper insights into how participants perceived the avatars’ races, we employed Principle Component Analysis (PCA) to reduce the agreement rates of each of the 42 base avatars down to two dimensions. Next, we performed K-means clustering [39] on the resulting two-dimensional data to group the avatars based on their perceived race. We optimized the number of clusters using the elbow method and distortion scores [5]. We applied this technique to both other-race and same-race agreement rates to determine whether there were any differences in

the clustering based on participant race. By visualizing the clusters, we aimed to better understand the differences in how participants perceived the avatars’ races.

Figure 4 shows that Asian, Black, and White avatars were perceived consistently by all participants, with clearly defined clusters. However, there was more confusion in perceiving AIAN, Hispanic, MENA, and NHPI avatars, which clustered closer together. Same-race participants had less overlap and more-accurately perceived these avatars, with more separation between them. For example, the Hispanic and MENA avatars were in separate clusters for same-race participants, except for one avatar (`Hispanic_F_2`). On the other hand, the Hispanic and MENA avatars were entirely clustered together for other-race participants.

## 6 DISCUSSION

In this section, we discuss the validation of our avatars. Specifically, we examine the extent to which each avatar was correctly identified as its intended race and the variability in identification across different participant groups. Additionally, we discuss the implications of our results for virtual avatar research, highlighting the importance of considering the potential impact of own-race bias on avatar race perception. Finally, we describe the potential future impact of our avatar library in the community, including how it can be used to promote diversity and inclusion.

### 6.1 Race Identification

**6.1.1 Universally Identified Avatars.** We found that our Asian, Black, and White avatars were recognized by all participants with high agreement rates. This suggests that these avatars can be a valuable tool for researchers seeking to create virtual humans that can be easily identified by individuals from different racial backgrounds.

Our results may be due to perceptual expertise or familiarity with other-race faces, as proposed by Civile et al. [14]. We hypothesize that this familiarity could be explained by the prevalence of these racial groups in global media and pop culture. For example, White cast members were the most represented in popular Hollywood movies over the last decade, followed by Black cast members [44]. Since Hollywood movies have a dominant share in the global film industry [43], people may be more familiar with characters that are prevalent in these films. Additionally, East Asian media culture has become widely popular worldwide over the past few decades [31, 33]. Phenomena like “The Hallyu Wave” and “Cool Japan” [40] have enabled East Asian films, dramas, and pop music to gain a global following. As people may often encounter these racial groups in media, this familiarity may have facilitated their recognition of these avatars.

**6.1.2 Same-Race Identified Avatars.** As expected, some avatars were only identified by participants of the same race as the avatar, consistent with the own-race bias effect. For example, as seen in Table 3, the Hispanic avatars received mixed ratings of White and Hispanic across all participants, but most were perceived as solely Hispanic by Hispanic-only participants. Similarly, only one MENA avatar was perceived as MENA by all participants, while five were perceived as MENA by MENA-only participants. These results suggest that participants’ own-race bias, a well-known phenomenon in psychology, may also affect their perception of virtual avatars. The



**Table 3: Assigned labels for all 42 base avatars. “All” indicates that the label was identified by all 132 participants, while “Same-Race” only includes the data of participants who identify as the race that the avatar was modeled for. Agreement labels were calculated as the percentage of participants who perceived an avatar to represent a race or gender.**

Avatar	Race (All)	Agreement	Race (Same-Race)	Agreement	Gender (All)	Agreement
AIAN_M_1	AIAN	0.67	AIAN	0.90	Male	0.89
AIAN_M_2	AIAN	0.65	AIAN	0.75	Male	0.89
AIAN_F_1	AIAN	0.55	AIAN	0.60	Female	0.64
AIAN_F_2	AIAN	0.55	AIAN	0.75	Female	0.75
Asian_M_1	Asian	0.86	Asian	0.90	Male	0.95
Asian_M_2	Asian	0.98	Asian	1.00	Male	0.91
Asian_M_3	Asian	0.98	Asian	1.00	Male	0.91
Asian_F_1	Asian	0.98	Asian	1.00	Female	0.98
Asian_F_2	Asian	0.93	Asian	0.90	Female	0.96
Asian_F_3	Asian	0.95	Asian	1.00	Female	0.97
Black_M_1	Black	0.98	Black	1.00	Male	0.98
Black_M_2	Black	0.97	Black	1.00	Male	0.98
Black_M_3	Black	0.98	Black	1.00	Male	0.73
Black_F_1	Black	0.98	Black	1.00	Female	0.98
Black_F_2	Black	0.98	Black	1.00	Female	0.94
Black_F_3	Black	0.95	Black	1.00	Female	0.97
Hispanic_M_1	Hispanic/White	0.65/0.55	Hispanic	0.85	Male	0.98
Hispanic_M_2	White	0.64	Hispanic	0.80	Male	0.98
Hispanic_M_3	Hispanic/White	0.55/0.51	Hispanic	0.75	Male	0.97
Hispanic_F_1	Hispanic	0.64	Hispanic	0.80	Female	0.95
Hispanic_F_2	White	0.70	Hispanic/White	0.55/0.60	Female	0.95
Hispanic_F_3	Hispanic/White	0.59/0.52	Hispanic	0.75	Female	0.93
MENA_M_1	White	0.56	MENA	0.70	Male	0.99
MENA_M_2	White	0.64	MENA/White	0.65/0.60	Male	1.00
MENA_M_3	MENA/White	0.55/0.65	MENA/White	0.75/0.55	Male	0.98
MENA_F_1	White	0.58	MENA/White	0.70/0.60	Female	0.98
MENA_F_2	White	0.60	MENA	0.60	Female	0.98
NHPI_M_1	NHPI	0.52	NHPI	0.58	Male	0.98
NHPI_M_2	Hispanic	0.65	NHPI	0.67	Male	1.00
White_M_1	White	0.96	White	0.95	Male	0.99
White_M_2	White	0.98	White	0.95	Male	1.00
White_M_3	White	0.93	White	0.90	Male	0.99
White_F_1	White	0.94	White	0.95	Female	0.97
White_F_2	White	0.96	White	0.95	Female	0.98
White_F_3	White	0.94	White	0.95	Female	0.99
X_AIAN_M_1	Hispanic	0.64	Hispanic	0.75	Male	0.99
X_AIAN_F_1	Hispanic	0.54	Hispanic	0.45	Female	0.86
X_MENA_F_1	Ambiguous	N/A	Ambiguous	N/A	Female	0.98
X_NHPI_M_1	Hispanic	0.55	Asian	0.58	Male	0.98
X_NHPI_F_1	Hispanic	0.55	Ambiguous	N/A	Female	0.92
X_NHPI_F_2	Hispanic	0.58	Ambiguous	N/A	Female	0.95
X_NHPI_F_3	NHPI	0.52	Ambiguous	N/A	Female	0.92

findings point to the importance of considering participants’ race when using virtual avatars in research or applications that require accurate representation of different racial groups.

*6.1.3 Ambiguous Avatars.* Several avatars in our library were perceived ambiguously by all participants and only same-race participants, and therefore labeled as such (see Table 3 for details). Identifying the reason for these avatars’ lack of clear identification is not straightforward, and multiple factors could be at play. For instance, the two ambiguous AIAN avatars were the only ones with

short hairstyles, which may have impacted their identification as AIAN. Long hair carries cultural and spiritual significance in many AIAN tribes [71], and some participants may have perceived the avatars as non-AIAN as a result, even among AIAN participants.

The validation of our NHPI avatars was limited, possibly due to the low number of NHPI participants ( $n = 12$ ) in our study, despite our targeted recruitment efforts. As a consequence, most of the NHPI avatars were not validated by NHPI participants, including the lack of validation for any female NHPI avatars. Another potential reason for this lack of validation is that the majority of our NHPI

Avatar	Other-Race Participants							Same-Race Participants						
	AIAN	Asian	Black	Hispanic	MENA	NHPI	White	AIAN	Asian	Black	Hispanic	MENA	NHPI	White
AIAN_M_1	0.63	0.04	0.09	0.40	0.11	0.46	0.12	0.90	0.05	0.05	0.20	0.15	0.10	0.05
AIAN_M_2	0.63	0.03	0.07	0.45	0.12	0.41	0.09	0.75	0.00	0.00	0.30	0.10	0.40	0.10
AIAN_F_1	0.54	0.08	0.04	0.48	0.20	0.35	0.21	0.60	0.00	0.00	0.40	0.10	0.25	0.15
AIAN_F_2	0.51	0.09	0.08	0.53	0.23	0.31	0.21	0.75	0.05	0.00	0.15	0.30	0.25	0.15
Asian_M_1	0.12	0.85	0.00	0.12	0.03	0.19	0.21	0.15	0.90	0.00	0.05	0.00	0.15	0.10
Asian_M_2	0.11	0.97	0.00	0.05	0.02	0.10	0.07	0.05	1.00	0.00	0.00	0.00	0.00	0.10
Asian_M_3	0.05	0.98	0.00	0.04	0.01	0.12	0.11	0.00	1.00	0.00	0.00	0.00	0.10	0.15
Asian_F_1	0.08	0.97	0.00	0.04	0.03	0.12	0.06	0.00	1.00	0.05	0.00	0.00	0.05	0.05
Asian_F_2	0.12	0.94	0.01	0.04	0.04	0.18	0.05	0.20	0.90	0.00	0.00	0.05	0.05	0.05
Asian_F_3	0.18	0.95	0.00	0.06	0.03	0.13	0.07	0.05	1.00	0.05	0.05	0.05	0.05	0.00
Black_M_1	0.02	0.04	0.97	0.05	0.15	0.12	0.02	0.00	0.00	1.00	0.00	0.25	0.05	0.00
Black_M_2	0.03	0.01	0.96	0.05	0.20	0.04	0.00	0.00	0.00	1.00	0.05	0.15	0.00	0.00
Black_M_3	0.03	0.02	0.98	0.04	0.18	0.04	0.00	0.00	0.00	0.95	0.00	0.25	0.00	0.00
Black_F_1	0.01	0.02	0.98	0.04	0.17	0.06	0.00	0.00	0.00	1.00	0.05	0.25	0.00	0.00
Black_F_2	0.03	0.02	0.97	0.05	0.12	0.07	0.00	0.00	0.00	1.00	0.05	0.25	0.00	0.00
Black_F_3	0.03	0.02	0.94	0.07	0.16	0.08	0.00	0.00	0.00	1.00	0.05	0.10	0.05	0.00
Hispanic_M_1	0.11	0.05	0.03	0.60	0.29	0.12	0.58	0.25	0.15	0.00	0.85	0.25	0.10	0.40
Hispanic_M_2	0.12	0.06	0.02	0.46	0.19	0.12	0.71	0.25	0.00	0.00	0.80	0.25	0.20	0.30
Hispanic_M_3	0.14	0.08	0.03	0.52	0.26	0.14	0.54	0.10	0.10	0.00	0.75	0.10	0.20	0.30
Hispanic_F_1	0.13	0.09	0.04	0.61	0.31	0.14	0.43	0.15	0.00	0.05	0.80	0.30	0.10	0.15
Hispanic_F_2	0.08	0.05	0.04	0.52	0.15	0.09	0.71	0.10	0.10	0.00	0.55	0.30	0.05	0.60
Hispanic_F_3	0.23	0.08	0.05	0.56	0.22	0.09	0.55	0.15	0.05	0.00	0.75	0.25	0.15	0.30
MENA_M_1	0.09	0.03	0.00	0.50	0.46	0.05	0.59	0.10	0.05	0.00	0.25	0.70	0.05	0.40
MENA_M_2	0.06	0.04	0.00	0.34	0.47	0.07	0.64	0.10	0.05	0.00	0.25	0.65	0.10	0.60
MENA_M_3	0.06	0.02	0.01	0.38	0.51	0.05	0.67	0.10	0.05	0.00	0.45	0.75	0.05	0.55
MENA_F_1	0.12	0.09	0.02	0.49	0.31	0.11	0.58	0.10	0.15	0.00	0.25	0.70	0.10	0.60
MENA_F_2	0.20	0.07	0.02	0.41	0.27	0.10	0.62	0.15	0.10	0.00	0.40	0.60	0.10	0.45
NHPI_M_1	0.28	0.34	0.07	0.45	0.16	0.52	0.08	0.25	0.42	0.00	0.33	0.00	0.58	0.00
NHPI_M_2	0.28	0.17	0.08	0.66	0.27	0.46	0.15	0.17	0.25	0.00	0.42	0.08	0.67	0.00
White_M_1	0.03	0.00	0.00	0.08	0.10	0.03	0.96	0.05	0.00	0.00	0.05	0.00	0.05	0.95
White_M_2	0.05	0.01	0.00	0.10	0.08	0.04	0.98	0.05	0.00	0.00	0.05	0.10	0.10	0.95
White_M_3	0.04	0.02	0.00	0.15	0.09	0.04	0.94	0.00	0.00	0.00	0.05	0.10	0.10	0.90
White_F_1	0.06	0.04	0.00	0.17	0.06	0.04	0.94	0.10	0.00	0.05	0.05	0.00	0.10	0.95
White_F_2	0.05	0.04	0.00	0.09	0.06	0.03	0.96	0.05	0.05	0.00	0.00	0.00	0.00	0.95
White_F_3	0.08	0.01	0.00	0.19	0.05	0.05	0.94	0.00	0.00	0.00	0.15	0.00	0.05	0.95
X_AIAN_M_1	0.22	0.16	0.04	0.62	0.29	0.25	0.33	0.15	0.10	0.10	0.75	0.15	0.25	0.35
X_AIAN_F_1	0.30	0.24	0.03	0.55	0.21	0.27	0.27	0.25	0.15	0.00	0.45	0.15	0.25	0.35
X_MENA_F_1	0.14	0.08	0.04	0.43	0.33	0.08	0.48	0.20	0.10	0.00	0.40	0.45	0.05	0.50
X_NHPI_M_1	0.33	0.25	0.03	0.60	0.23	0.56	0.11	0.25	0.58	0.00	0.33	0.00	0.42	0.00
X_NHPI_F_1	0.29	0.24	0.34	0.56	0.20	0.42	0.05	0.17	0.17	0.25	0.50	0.08	0.42	0.25
X_NHPI_F_2	0.30	0.19	0.30	0.59	0.28	0.38	0.08	0.25	0.00	0.33	0.42	0.25	0.50	0.08
X_NHPI_F_3	0.33	0.33	0.14	0.49	0.17	0.54	0.05	0.42	0.08	0.00	0.42	0.00	0.42	0.00

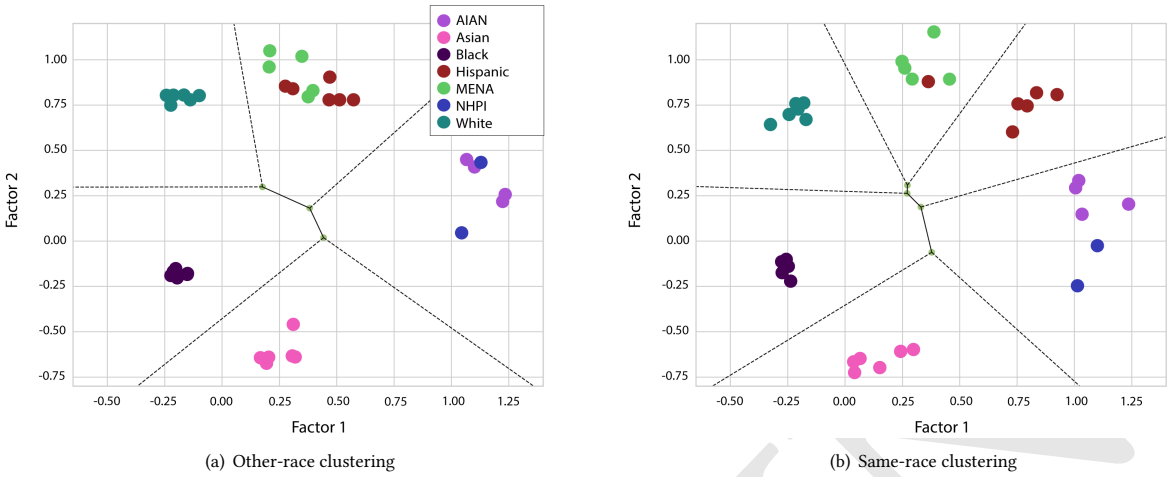
**Figure 3: Confusion matrix heatmap of agreement rates for the 42 base avatars by separated by other-race participants and same-race participants (i.e., participants of a different or same race as the avatar). Agreement rates were calculated as the percentage of participants who perceived an avatar to represent a race or gender.**

participants identified themselves as New Zealand Maori, whereas our avatars were developed with the help of Samoan and Native Hawaiian volunteer representatives. Therefore, it is possible that our NHPI avatars are representative of some NHPI cultures, but not New Zealand Maori. In future studies, expanding recruitment efforts for both interview volunteers and study participants will be crucial, despite the challenges involved in doing so. For example, future studies may need to compensate NHPI participants more than participants of other races.

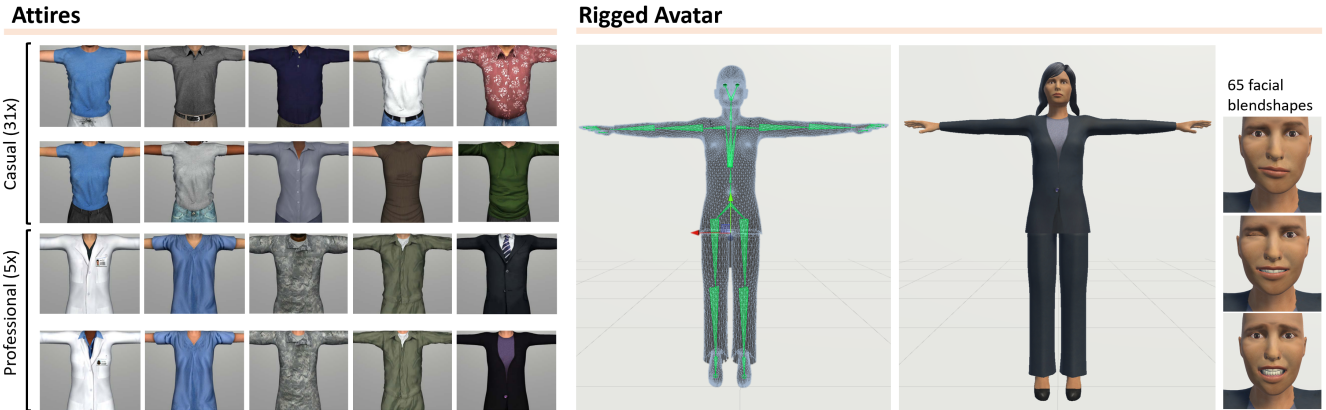
## 6.2 Implications for Virtual Avatars

Our study provides valuable insights for virtual avatar applications and research. Our findings indicate that human behavior in race categorization can apply to virtual avatars, which has notable implications for interactions in virtual experiences. Kawakami et al. [36] suggest that in-group and out-group categorization can lead to stereotyping, social judgments, and group-based evaluations. Therefore, designers and developers should be aware of this and take necessary steps to mitigate unintended consequences in virtual experiences. For example, regulating codes of conduct [70] can help to improve interracial interactions in VR.





**Figure 4: Clustered scatterplots of each avatar’s relation to one another based on Principle Component Analysis and K-means clustering for other-race and same-race participant identifications. The Voronoi analysis shows the borders of the clusters where each category was assigned. Each avatar is color coded by its validated label.**



**Figure 5: Images of the skeleton and facial blend shapes included with our avatars.**

Interestingly, our study also replicated a nuanced finding from more recent psychology research on the perception of ambiguous avatars [51]. As seen in Table 3, most of the misidentified avatars were identified as Hispanic by all participants. Similarly, Nicolas et al. [51] recently found that participants classify racially ambiguous photos as Hispanic or MENA, regardless of their parent ethnicities. We believe that this effect extended to our virtual avatars.

### 6.3 An Open Library of Validated Avatars

As a contribution to the research community, we are providing open access to our virtual avatar library, which includes all 210 fully rigged avatars, along with validated labels for each avatar’s race and gender. Our library features avatars of seven different races, providing a diverse selection for researchers to use in their studies. The validated labels can facilitate research on the impact of

avatar race, and researchers can choose to use the labels for studies aimed at individuals from different racial backgrounds or same-race labels for specific study populations.

The *Virtual Avatar Library for Inclusion and Diversity (VALID)* provides researchers and developers with a diverse set of fully rigged avatars suitable for various scenarios such as casual, business, medical, military, and utility. Each avatar comes with 65 facial blend shapes, enabling dynamic facial expressions (see Figure 5). The library is readily available for download and can be used in popular game engines like Unity or Unreal. Although this is the first iteration of the library, we plan to update it by adding more professions and outfits soon. In addition, the library can be used for a wide range of research purposes, including social psychology simulations and educational applications.

## 6.4 Limitations and Future Work

We recognize that our VALID avatar library is only a small step towards achieving greater diversity and inclusion in avatar resources. We acknowledge that the representation of each demographic is limited and plan to expand the diversity within each group by creating new avatars. For example, our Asian avatars are modeled after East Asians, but we plan to expand VALID to include South Asian and Southeast Asian avatars as well. Our Hispanic representatives have pointed out the need for more diverse Hispanic avatars, including varying skin tones to represent different South American populations, such as Mexican and Cuban. Additionally, our NHPI representatives have suggested the inclusion of tattoos, which hold cultural significance for some NHPI communities, could improve the identifiability of our NHPI avatars, in addition to improving our NHPI recruitment methods. Any future updates to the library will undergo the same rigorous creation, iteration, and validation process as the current avatars.

While our first iteration of the library focused on diversity in terms of race, we realize that the avatars mostly represent young and fit adults, which does not reflect all types of people. In the future, we plan to update the library with a diversity of body types that include different body mass index (BMI) representations and ages. Including avatars with different BMI representations is not only more inclusive, but can also be useful for studies targeting physical activity, food regulation, and therapy [65]. Likewise, we plan to include shaders and bump maps [30] that can age any given avatar by creating realistic wrinkles and skin folds, further improving the diversity and inclusivity of VALID

Another limitation of the current work is that our library includes only male and female representations. In future updates, we plan to include non-binary and androgynous avatars. Currently, there are not many androgynous models that are freely available. However, they can be an area of important study. For example, previous studies found that androgynous avatars reduce gender bias and stereotypical assumptions in virtual agents [50] and improve student attitudes [10]. Thus, we plan to include these avatars in a future update by following Nag et al.'s [50] guidelines for creating androgynous virtual humans.

Our study, while diverse in terms of race and country, is not representative of everyone. We recruited participants through the online platform Prolific, which is known for its increased diversity compared to other crowdsourcing platforms such as Mechanical Turk. However, due to the online nature of the platform, we primarily recruited younger adults. It is possible that perceptions of our avatars may differ among other age groups, such as children or older adults. Therefore, it is important to broaden recruitment efforts by exploring alternative platforms and recruitment strategies that may be more effective in reaching a wider range of participants. Future studies could also consider conducting in-person studies or focus groups to gather additional insights into avatar perception.

## 7 CONCLUSION

We have introduced a new virtual avatar library comprised of 210 fully rigged avatars with diverse professions and outfits, available

for free. Our library aims to promote diversity and inclusion by creating equitable representation of seven races across various professions. We designed 42 base avatars using data-driven facial averages and collaborated with volunteer representatives of each ethnicity. A large validation study involving participants from around the world was conducted to obtain validated labels and metadata for the perceived race and gender of each avatar. Additionally, we offer a comprehensive process for creating, iterating, and validating diverse avatars to aid other researchers in creating similarly validated avatars.

Our validation study revealed that the majority of avatars were accurately perceived as the race they were modeled for. However, we observed that some avatars, such as the Hispanic and MENA avatars, were only validated as such by participants who identified as Hispanic or MENA, respectively. This finding suggests that the perception of virtual avatars may be influenced by own-race bias or the other-race effect, as described in the psychology literature. Moving forward, we plan to expand the library to include additional races, professions, body types, age ranges, and gender representations to further improve diversity and inclusion.

## REFERENCES

- [1] Mohammad Rafayet Ali, Seyede Zahra Razavi, Raina Langevin, Abdullah Al Mamun, Benjamin Kane, Reza Rawassizadeh, Lenhart K. Schubert, and Ehsan Hoque. 2020. A Virtual Conversational Agent for Teens with Autism Spectrum Disorder: Experimental Results and Design Lessons. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents (Virtual Event, Scotland, UK) (IVA '20)*. Association for Computing Machinery, New York, NY, USA, Article 2, 8 pages. <https://doi.org/10.1145/3383652.3423900>
- [2] Wilma A. Bainbridge, Phillip Isola, and Aude Oliva. 2013. The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General* 142, 4 (2013), 1323–1334. <https://doi.org/10.1037/a0033872>
- [3] Amy Baylor. 2003. Effects of Images and Animation on Agent Persona. *Journal of Educational Computing Research* 28, 4 (2003), 373–394.
- [4] Amy L. Baylor. 2009. Promoting motivation with virtual agents and avatars: Role of visual presence and appearance. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1535 (2009), 3559–3565. <https://doi.org/10.1098/rstb.2009.0148>
- [5] Benjamin Bengfort, Rebecca Bilbro, Nathan Danielsen, Larry Gray, Kristen McIntyre, Prema Roman, Zijie Poh, et al. 2018. *Yellowbrick*. <https://doi.org/10.5281/zenodo.1206264>
- [6] Timothy Bickmore, Dhaval Parmar, Everlyne Kimani, and Stefan Olafsson. 2021. Diversity Informatics: Reducing Racial and Gender Bias with Virtual Agents. In *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents (Virtual Event, Japan) (IVA '21)*. Association for Computing Machinery, New York, NY, USA, 25–32. <https://doi.org/10.1145/3472306.3478365>
- [7] Jim Blascovich, Natalie A. Wyer, Laura A. Swart, and Jeffrey L. Kibler. 1997. Racism and racial categorization. *Journal of Personality and Social Psychology* 72, 6 (1997), 1364–1372. <https://doi.org/10.1037/0022-3514.72.6.1364>
- [8] Kristopher J. Blom, Anna I. Bellido Rivas, Xenxo Alvarez, Ozan Cetinaslan, Bruno Oliveira, Verónica Orvalho, and Mel Slater. 2014. Achieving participant acceptance of their avatars. *Presence: Teleoperators and Virtual Environments* 23, 3 (2014), 287–299.
- [9] Marion Boberg, Petri Piippo, and Elina Ollila. 2008. Designing avatars. In *Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts*. ACM, Athens Greece, 232–239. <https://doi.org/10.1145/1413634.1413679>
- [10] Shirley Booth, Sara Goodman, Gill Kirkup, Agneta Gulz, and Magnus Haake. 2010. *Challenging Gender Stereotypes Using Virtual Pedagogical Characters*. 113–132. <https://doi.org/10.4018/978-1-61520-813-5.ch007>
- [11] Crystal Butler, Stephanie Michalowicz, Lakshmi Subramanian, and Winslow Burleson. 2017. More than a Feeling: The MiFace Framework for Defining Facial Communication Mappings. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (Québec City, QC, Canada) (UIST '17)*. Association for Computing Machinery, New York, NY, USA, 773–786. <https://doi.org/10.1145/3126594.3126640>
- [12] Elizabeth N. Chapman, Anna Kaatz, and Molly Carnes. 2013. Physicians and implicit bias: how doctors may unwittingly perpetuate health care disparities. *Journal of general internal medicine* 28, 11 (2013), 1504–1510.

- [13] Yi Fei Cheng, Hang Yin, Yukang Yan, Jan Gugenheimer, and David Lindlbauer. 2022. Towards Understanding Diminished Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 549, 16 pages. <https://doi.org/10.1145/3491102.3517452>
- [14] Ciro Civile and I. P. L. McLaren. 2022. Transcranial direct current stimulation (tDCS) eliminates the other-race effect (ORE) indexed by the face inversion effect for own versus other-race faces. *Scientific Reports* 12, 1 (2022), 1–10. <https://doi.org/10.1038/s41598-022-17294-w>
- [15] Lisa DeBruine. 2018. *debruine/webmorph: Beta release 2*. <https://doi.org/10.5281/zenodo.1162670>
- [16] Diane Dewez, Ludovic Hoyet, Anatole Lécuyer, and Ferran Argelaguet Sanz. 2021. Towards “Avatar-Friendly” 3D Manipulation Techniques: Bridging the Gap Between Sense of Embodiment and Interaction in Virtual Reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 264, 14 pages. <https://doi.org/10.1145/3411764.3445379>
- [17] Tiffany D. Do, Ryan P. McMahan, and Pamela J. Wisniewski. 2022. A New Uncanny Valley? The Effects of Speech Fidelity and Human Listener Gender on Social Perceptions of a Virtual-Human Speaker. In *CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 424, 11 pages. <https://doi.org/10.1145/3491102.3517564>
- [18] Jesse Fox, Sun Joo (Grace) Ahn, Joris H. Janssen, Leo Yeykelis, Kathryn Y. Segovia, and Jeremy N. Bailenson. 2015. Avatars Versus Agents: A Meta-Analysis Quantifying the Effect of Agency on Social Influence. *Human-Computer Interaction* 30, 5 (2015), 401–432. <https://doi.org/10.1080/07370024.2014.921494> arXiv:<https://doi.org/10.1080/07370024.2014.921494>
- [19] Luciano Gamberini, Luca Chittaro, Anna Spagnoli, and Claudio Carlesso. 2015. Psychological response to an emergency in virtual reality: Effects of victim ethnicity and emergency type on helping behavior and navigation. *Computers in Human Behavior* 48 (2015), 104–113. <https://doi.org/10.1016/j.chb.2015.01.040>
- [20] Liezhong Ge, Hongchuan Zhang, Zhe Wang, Paul C Quinn, Olivier Pascalis, David Kelly, Alan Slater, Jie Tian, and Kang Lee. 2009. Two faces of the other-race effect: Recognition and categorisation of Caucasian and Chinese faces. *Perception* 38, 8 (2009), 1199–1210. <https://doi.org/10.1068/p6136>
- [21] Juan E Gilbert, Keena Arbuthnot, Stafford Hood, Michael M Grant, Melanie L West, Yolanda Mcmillian, E Vincent Cross, Philiicity Williams, and Wanda Eugene. 2008. Teaching Algebra Using Culturally Relevant Virtual Instructors. *The International Journal of Virtual Reality* 7, 1 (2008), 21–30.
- [22] Mar Gonzalez-Franco, Anna I Bellido, Kristopher J Blom, Mel Slater, and Antoni Rodriguez-Fornells. 2016. The neurological traces of look-alike avatars. *Frontiers in human neuroscience* 10 (2016), 392.
- [23] Mar Gonzalez-Franco, Zelia Egan, Matthew Peachey, Angus Antley, Tanmay Randhavane, Payod Panda, Yaying Zhang, Cheng Yao Wang, Derek F. Reilly, Tabitha C Peck, Andrea Stevenson Won, Anthony Steed, and Eyal Ofek. 2020. MoveBox: Democratizing MoCap for the Microsoft Rocketbox Avatar Library. In *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. 91–98. <https://doi.org/10.1109/AIVR50618.2020.00026>
- [24] Mar Gonzalez-Franco, Eyal Ofek, Ye Pan, Angus Antley, Anthony Steed, Bernhard Spanlang, Antonella Maselli, Domna Banakou, Nuria Pelechano, Sergio Orts-Escolano, Veronica Orvalho, Laura Trutoiu, Markus Wojcik, Maria V. Sanchez-Vives, Jeremy Bailenson, Mel Slater, and Jaron Lanier. 2020. The Rocketbox Library and the Utility of Freely Available Rigged Avatars. *Frontiers in Virtual Reality* 1 (2020). <https://doi.org/10.3389/frvir.2020.561558>
- [25] Mar Gonzalez-Franco, Mel Slater, Megan E Birney, David Swapp, S Alexander Haslam, and Stephen D Reicher. 2018. Participant concerns for the Learner in a Virtual Reality replication of the Milgram obedience study. *PLoS one* 13, 12 (2018), e0209704.
- [26] Victoria Groom, Jeremy N Bailenson, and Clifford Nass. 2009. The influence of racial embodiment on racial bias in immersive virtual environments. *Social Influence* 4, 3 (2009), 231–248.
- [27] Shivashankar Halan, Isaac Sia, Michael Cray, and Benjamin Lok. 2015. Exploring the effects of healthcare students creating virtual patients for empathy training. In *Intelligent Virtual Agents: 15th International Conference, IVA 2015, Delft, The Netherlands, August 26-28, 2015, Proceedings 15*. Springer, 239–249.
- [28] Paul Heidicker, Eike Langbehn, and Frank Steinicke. 2017. Influence of avatar appearance on presence in social VR. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. 233–234. <https://doi.org/10.1109/3DUI.2017.7893357>
- [29] Ann Huang, Pascal Knierim, Francesco Chiassi, Lewis L Chuang, and Robin Welsch. 2022. Proxemics for Human-Agent Interaction in Augmented Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 421, 13 pages. <https://doi.org/10.1145/3491102.3517593>
- [30] Rowan Hughes. 2011. *Real-Time Wrinkles For Expressive Virtual Characters*. Ph.D. Dissertation. Trinity College Dublin.
- [31] Koichi Iwabuchi. 2010. Globalization, East Asian media cultures and their publics. *Asian Journal of Communication* 20, 2 (2010), 197–212. <https://doi.org/10.1080/01292981003693385>
- [32] David C. Jeong, Jackie Jingyi Xu, and Lynn C. Miller. 2020. Inverse Kinematics and Temporal Convolutional Networks for Sequential Pose Analysis in VR. In *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. 274–281. <https://doi.org/10.1109/AIVR50618.2020.00056>
- [33] D.Y. Jin, K. Yoon, and W. Min. 2021. *Transnational Hallyu: The Globalization of Korean Digital and Popular Culture*. Rowman & Littlefield Publishers. <https://books.google.com/books?id=zA8zEAAQBAJ>
- [34] Dongsik Jo, Kangsoo Kim, Gregory F. Welch, Woojin Jeon, Yongwan Kim, Ki-Hong Kim, and Gerard Jounghyun Kim. 2017. 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) (VRST '17). Association for Computing Machinery, New York, NY, USA, Article 77, 2 pages. <https://doi.org/10.1145/3139131.3141214>
- [35] Sungchul Jung, Gerd Bruder, Pamela J. Wisniewski, Christian Sandor, and Charles E. Hughes. 2018. Over My Hand: Using a Personalized Hand in VR to Improve Object Size Estimation, Body Ownership, and Presence. In *Proceedings of the Symposium on Spatial User Interaction* (Berlin, Germany) (SUI '18). Association for Computing Machinery, New York, NY, USA, 60–68. <https://doi.org/10.1145/3267782.3267920>
- [36] K. Kawakami, D.M. Amodio, and K. Hugenberg. 2017. Chapter One - Intergroup Perception and Cognition: An Integrative Framework for Understanding the Causes and Consequences of Social Categorization. *Advances in Experimental Social Psychology*, Vol. 55. Academic Press, 1–80. <https://doi.org/10.1016/b.s.aesp.2016.10.001>
- [37] Yanghee Kim and Jae Hoon Lim. 2013. Gendered Socialization with an Embodied Agent: Creating a Social and Affable Mathematics Learning Environment for Middle-Grade Females. *Journal of Educational Psychology* 105, 4 (2013), 1164–1174. <https://doi.org/10.1037/a0031027>
- [38] Jiajia Li, Zixia Zheng, Xiemin Wei, and Guanyun Wang. 2021. FaceMe: An Augmented Reality Social Agent Game for Facilitating Children’s Learning about Emotional Expressions. In *Adjunct Proceedings of the 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) (UIST '21 Adjunct). Association for Computing Machinery, New York, NY, USA, 17–19. <https://doi.org/10.1145/3474349.3480216>
- [39] Aristidis Likas, Nikos Vlassis, and Jakob J. Verbeek. 2003. The global k-means clustering algorithm. *Pattern Recognition* 36, 2 (2003), 451–461. [https://doi.org/10.1016/S0031-3203\(02\)00060-2](https://doi.org/10.1016/S0031-3203(02)00060-2) Biometrics.
- [40] Gillianne Lux. 2021. Cool Japan and the Hallyu Wave : The Effect of Popular Culture Exports on National Image and Soft Power. *East Asian Studies Honors Paper* (2021). [https://digitalcommons.ursinus.edu/eastasia\\_hon/3%0AThis](https://digitalcommons.ursinus.edu/eastasia_hon/3%0AThis)
- [41] Otto H. MacLin and Roy S. Malpass. 2001. Racial Categorization of Faces: The Ambiguous Race Face Effect. *Psychology, Public Policy, and Law* 7, 1 (2001), 98–118. <https://doi.org/10.1037/1076-8971.7.1.98>
- [42] Lara Maister, Mel Slater, Maria V. Sanchez-Vives, and Manos Tsakiris. 2015. Changing bodies changes minds: Owning another body affects social cognition. *Trends in Cognitive Sciences* 19, 1 (2015), 6–12. <https://doi.org/10.1016/j.tics.2014.11.001>
- [43] Wanwarang Maisuwong. 2012. The Promotion of American Culture Through Hollywood Movies to The World. *International Journal of Engineering Research and Technology (IJERT)* Vol. 1, 4, June (2012), 1–7. <https://www.ijert.org/research/the-promotion-of-american-culture-through-hollywood-movies-to-the-world-IJERTV1IS4194.pdf>
- [44] Musa Malik, Frederic R. Hopp, and René Weber. 2022. Representations of Racial Minorities in Popular Movies. *Computational Communication Research* 4, 1 (2022), 208–253. <https://doi.org/10.5117/ccr2022.1.006.mali>
- [45] K. Matthews, J. Phelan, N. A. Jones, S. Konya, R. Marks, B. M. Pratt, J. Combs, and M. Bentley. 2017. *2015 National Content Test Race and Ethnicity Analysis Report*. Technical Report. U.S. Department of Commerce. 1–380 pages. <https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2015nct-race-ethnicity-analysis.pdf>
- [46] Jess McIntosh, Hubert Dariusz Zajac, Andreea Nicoleta Stefan, Joanna Bergström, and Kasper Hornbæk. 2020. Iteratively Adapting Avatars Using Task-Integrated Optimisation. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) (UIST '20). Association for Computing Machinery, New York, NY, USA, 709–721. <https://doi.org/10.1145/3379337.3415832>
- [47] Christian A Meissner and John C Brigham. 2001. Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law* 7, 1 (2001), 3.
- [48] Roxana Moreno and Terri Flowerday. 2006. Students’ choice of animated pedagogical agents in science learning: A test of the similarity-attraction hypothesis on gender and ethnicity. *Contemporary Educational Psychology* 31, 2 (2006), 186–207. <https://doi.org/10.1016/j.cedpsych.2005.05.002>
- [49] Masahiro Mori, Karl F. MacDorman, and Norri Kageki. 2012. The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine* 19, 2 (2012), 98–100. <https://doi.org/10.1109/MRA.2012.2192811>

- [50] Procheta Nag and Özge Nilay Yalçın. 2020. Gender Stereotypes in Virtual Agents. *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, IVA 2020* (2020). <https://doi.org/10.1145/3383652.3423876>
- [51] Gandalf Nicolas, Allison L. Skinner, and Cheryl L. Dickter. 2019. Other Than the Sum: Hispanic and Middle Eastern Categorizations of Black–White Mixed-Race Faces. *Social Psychological and Personality Science* 10, 4 (2019), 532–541. <https://doi.org/10.1177/1948550618769591>
- [52] Kristine L. Nowak and Frank Biocca. 2003. The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments* 12, 5 (10 2003), 481–494. <https://doi.org/10.1162/105474603322761289> arXiv:<https://direct.mit.edu/pvar/article-pdf/12/5/481/1623959/105474603322761289.pdf>
- [53] Nami Ogawa, Takuji Narumi, Hideaki Kuzuoka, and Michitaka Hirose. 2020. Do You Feel Like Passing Through Walls?: Effect of Self-Avatar Appearance on Facilitating Realistic Behavior in Virtual Environments. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376562>
- [54] Stefan Palan and Christian Schitter. 2018. Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance* 17 (2018), 22–27. <https://doi.org/10.1016/j.jbef.2017.12.004>
- [55] Tabitha C. Peck and Mar Gonzalez-Franco. 2021. Avatar Embodiment. A Standardized Questionnaire. *Frontiers in Virtual Reality* 1, February (2021), 1–12. <https://doi.org/10.3389/frvr.2020.575943>
- [56] Tabitha C. Peck, Jessica J. Good, and Katharina Seitz. 2021. Evidence of Racial Bias Using Immersive Virtual Reality: Analysis of Head and Hand Motions During Shooting Decisions. *IEEE Transactions on Visualization and Computer Graphics* 27, 5 (2021), 2502–2512. <https://doi.org/10.1109/TVCG.2021.3067767>
- [57] Tabitha C. Peck, Sofia Seinfeld, Salvatore M. Aglioti, and Mel Slater. 2013. Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition* 22, 3 (2013), 779–787. <https://doi.org/10.1016/j.concog.2013.04.016>
- [58] Eyal Peer, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. 2017. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology* 70 (2017), 153–163. <https://doi.org/10.1016/j.jesp.2017.01.006>
- [59] Catherine Pelachaud. 2009. Modelling multimodal expression of emotion in a virtual agent. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1535 (2009), 3539–3548. <https://doi.org/10.1098/rstb.2009.0186> arXiv:<https://royalsocietypublishing.org/doi/pdf/10.1098/rstb.2009.0186>
- [60] Thammathip Piumsomboon, Gun A. Lee, Jonathon D. Hart, Barrett Ens, Robert W. Lindeman, Bruce H. Thomas, and Mark Billinghurst. 2018. Mini-Me: An Adaptive Avatar for Mixed Reality Remote Collaboration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3173574.3173620>
- [61] Gillian Rhodes, Hanne C Lie, Louise Ewing, Emma Evangelista, and James W Tanaka. 2010. Does Perceived Race Affect Discrimination and Recognition of Ambiguous-Race Faces? A Test of the Sociocognitive Hypothesis. *Journal of Experimental Psychology: Learning Memory and Cognition* 36, 1 (2010), 217–223. <https://doi.org/10.1037/a0017680>
- [62] Gillian Rhodes, Vance Locke, Louise Ewing, and Emma Evangelista. 2009. Race coding and the other-race effect in face recognition. *Perception* 38, 2 (2009), 232–241. <https://doi.org/10.1068/p6110>
- [63] Natalie Salmanowitz. 2018. The impact of virtual reality on implicit racial bias and mock legal decisions. *Journal of Law and the Biosciences* 5, 1 (2018), 174–203. <https://doi.org/10.1093/jlb/lsy005>
- [64] Ayako Saneyoshi, Matia Okubo, Hikaru Suzuki, Takato Oyama, and Bruno Laeng. 2022. The other-race effect in the uncanny valley. *International Journal of Human Computer Studies* 166, March (2022), 102871. <https://doi.org/10.1016/j.ijhcs.2022.102871>
- [65] Federica Scarpina, Silvia Serino, Anouk Keizer, Alice Chirico, Massimo Scacchi, Gianluca Castelnuovo, Alessandro Mauro, and Giuseppe Riva. 2019. The effect of a virtual-reality full-body illusion on body representation in obesity. *Journal of Clinical Medicine* 8, 9 (2019). <https://doi.org/10.3390/jcm8091330>
- [66] Peipei Setoh, Kristy J.J. Lee, Lijun Zhang, Miao K Qian, Paul C Quinn, Gail D Heyman, and Kang Lee. 2019. Racial Categorization Predicts Implicit Racial Bias in Preschool Children. *Child Development* 90, 1 (2019), 162–179. <https://doi.org/10.1111/cdev.12851>
- [67] Ari Shapiro, Andrew Feng, Ruizhe Wang, Hao Li, Mark Bolas, Gerard Medioni, and Evan Suma. 2014. Rapid avatar capture and simulation using commodity depth sensors. *Computer Animation and Virtual Worlds* 25, 3-4 (2014), 201–211. <https://doi.org/10.1002/cav.1579> arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1002/cav.1579>
- [68] David J. Sheskin. 2011. *Handbook of Parametric and Nonparametric Statistical Procedures* (5th ed.). Chapman and Hall/CRC, New York, 18 pages. <https://doi.org/10.1201/9780429186196>
- [69] Misha Sra and Chris Schmandt. 2015. MetaSpace: Full-Body Tracking for Immersive Multiperson Virtual Reality. In *Adjunct Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Daegu, Kyungpook, Republic of Korea) (UIST '15 Adjunct). Association for Computing Machinery, New York, NY, USA, 47–48. <https://doi.org/10.1145/2815585.2817802>
- [70] Valerie Jones Taylor, Juan José Valladares, Claire Siepser, and Caitlyn Yantis. 2020. Interracial Contact in Virtual Reality: Best Practices. *Policy Insights from the Behavioral and Brain Sciences* 7, 2 (2020), 132–140. <https://doi.org/10.1177/2372732220943638> arXiv:<https://doi.org/10.1177/2372732220943638>
- [71] Saradadevi Thanikachalam, Nelson Sanchez, and Austin J. Maddy. 2019. Long Hair throughout the Ages. *Dermatology* 235, 3 (2019), 260–262. <https://doi.org/10.1159/000497156>
- [72] Matias Volonte, Eyal Ofek, Ken Jakubzak, Shawn Bruner, and Mar Gonzalez-Franco. 2022. HeadBox: A Facial Blendshape Animation Toolkit for the Microsoft Rocketbox Library. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 39–42. <https://doi.org/10.1109/VRW55335.2022.00015>
- [73] Thomas Waltemate, Dominik Gall, Daniel Roth, Mario Botsch, and Marc Erich Latoschik. 2018. The Impact of Avatar Personalization and Immersion on Virtual Body Ownership, Presence, and Emotional Response. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (2018), 1643–1652. <https://doi.org/10.1109/TVCG.2018.2794629>
- [74] Nick Yee and Jeremy Bailenson. 2007. The Proteus Effect: The Effect of Transformed Self-Representation on Behavior. *Human Communication Research* 33, 3 (07 2007), 271–290. <https://doi.org/10.1111/j.1468-2958.2007.00299.x> arXiv:<https://academic.oup.com/hcr/article-pdf/33/3/271/22324746/jhumcom0271.pdf>
- [75] Sarah A Zipp and Scotty D Craig. 2019. The impact of a user's biases on interactions with virtual humans and learning during virtual emergency management training. *Educational Technology Research and Development* 67, 6 (2019), 1385–1404.