

The Influence of Mixed-Gender Avatar Facial Features on Racial Perception: Insights from the VALID Avatar Library

Tiffany D. Do *
University of Central Florida

Camille Isabella Protko †
University of Central Florida

Ryan P. McMahan ‡
University of Central Florida

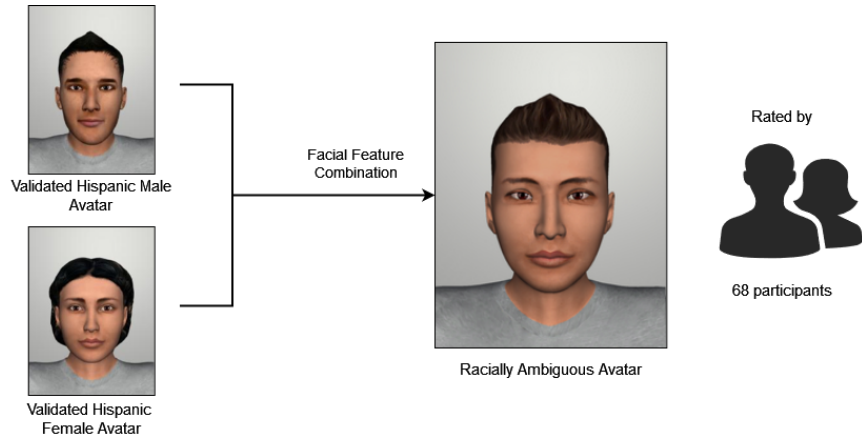


Figure 1: Two source avatars, perceptually categorized as Hispanic, were used to create a mixed-gender avatar by blending their facial features. The resulting mixed-gender avatar was consistently rated as racially ambiguous.

ABSTRACT

Using the VALID Avatar Library, we created 14 mixed-gender avatars by combining facial features from male and female avatars within seven racial categories. In an online study ($n=68$), the avatars were perceived as racially ambiguous, despite source avatars sharing the same race. Same-race participant analysis showed nuanced ambiguity reduction, emphasizing categorization refinement. Findings suggest a potential link between gender and race perception, influenced by limited perceptual expertise.

Index Terms: Human-centered computing—User studies; Computing methodologies—Perception

1 INTRODUCTION

The use of virtual reality (VR) has seen a notable rise in recent times, with many VR applications now incorporating avatars as representations within the virtual space as both self-avatars and simulated characters. As our understanding of gender diversity expands, it is crucial to create inclusive spaces, both physical and virtual, that cater to individuals identifying or presenting as gender diverse. Notably, younger demographics are more likely to design more androgynous avatars [9], and it is likely that these types of characters will be present in the future. However, there is limited research on how gender-diverse avatars are perceived, a factor that can profoundly influence users. For example, studies have shown that androgynous (i.e., gender neutral) avatars in virtual environments can challenge gender bias and stereotypes during interactions [6]. Additionally, androgynous avatars have demonstrated a positive impact on student

attitudes [1]. These findings emphasize the necessity for further exploration of the perception of gender-diverse virtual avatars. Despite its potential influence, research on androgynous virtual agents is still sparse. Furthermore, all virtual agents in these previously mentioned works were White, and there exists very little work investigating diverse androgynous or gender diverse avatars.

To expand on this research trajectory, using guidelines for creating androgynous avatars set forth by Nag and Yalcin [6], we developed 14 avatars with mixed-gender facial features by combining male and female avatars from the *Virtual Avatar Library for Inclusion and Diversity (VALID)* [3]. We created avatars of seven distinct racial categories: American Indian or Native Alaskan (AIAN), 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. We then conducted an online study with 68 participants to explore how mixing facial features from male and female avatars influences perceptions of race. Specifically, we sought to understand the potential interaction between gender perception and race perception.

Our investigation unveiled an unexpected finding: when combining facial attributes from male and female virtual avatars of identical racial backgrounds, avatars in the Hispanic, MENA, NHPI, and AIAN groups became considerably challenging to racially identify, despite their source avatars' validation within the same racial group. For instance, the fusion of facial traits from a male and a female MENA avatar resulted in an avatar presenting racial ambiguity, highlighting a relationship between gendered facial cues and the racial perception of virtual avatars. Additionally, we found that these avatars proved difficult to categorize even for participants of the same race as their source avatars, unlike the high accuracy observed in previous studies for categorizing male and female avatars of the same race [3].

These findings suggest a potential relationship between gendered facial cues and race perception for virtual avatars. Our study highlights the importance of caution in future research involving mixed-gender avatars, emphasizing the need for meticulous consideration when designing racially diverse, mixed-gender avatars

*e-mail: tiffany.do@ucf.edu

†e-mail: camille.protko@ucf.edu

‡e-mail: rpm@ucf.edu

2 RELATED WORK

2.1 VALID Avatar Library

We selected our source avatars from the *Virtual Avatar Library for Inclusion and Diversity (VALID)* [3], leveraging its diverse avatar representation and validated labels. These avatars were developed from facial averages of their respective ethnicities within the library and were perceptually validated in a large user study. Thus, we were able to mix facial attributes from both male and female avatars, ensuring that both avatars were perceived as the same race.

2.2 Perception of Androgynous Virtual Agents

Early research by Niculescu et al. [8] highlighted the challenge of understanding androgynous agents; participants tended to categorize agents as either feminine or masculine based on their own gender and the sequence of agent presentation. In subsequent work, Niculescu et al. [7] found that an androgynous Q&A virtual agent received lower ratings compared to a female agent. However, a more recent study by Nag and Yalcin [6] introduced specific guidelines for crafting a perceptually androgynous avatar, resulting in participants perceiving their virtual agents as gender-neutral.

In these guidelines, they note that combining certain facial cues from preset avatars could result in a gender neutral avatar. In a follow-up study, this androgynous agent served as a middle ground, challenging prevailing gender stereotypes and potentially assisting in alleviating these biases in future studies. Notably, their findings indicated a diminished impact of gender stereotypes compared to earlier studies. This reduced influence, coupled with enhanced 3D models, may explain the divergence in results from prior research. Correspondingly, Schwind et al. [9] observed that younger demographics design more androgynous avatars, indicating a growing acceptance of androgynous avatar designs, reflecting an increasing openness among future populations toward notions of androgynous and gender-diverse agents.

In a recent exploratory design study, Ghosh et al. [4] found that when undergraduate computing students were tasked with designing an androgynous avatar, they used inconsistent stereotypical face gender cues (i.e., sometimes the facial cue would be masculine and other times feminine). Additionally, they reported that face gender cues in general were not leveraged equally in creating an androgynous perception. These findings complement the guidelines proposed by Nag and Yalcin [6], shedding light on the nuanced nature of combining male and female facial features to create androgynous avatars. Together, these results suggest that while it is possible to mix masculine and feminine facial features to create androgynous avatars, the specific combination of features may differ between different avatars.

3 CREATION OF MIXED-GENDER AVATARS

Using the VALID library, which encompasses validated avatars for both race and gender, we incorporated distinct facial attributes from specific male and female avatars. It is important to note that our study exclusively selected avatars that were perceptually validated as their intended race. This decision was made since VALID also includes avatars that were not validated in terms of race.

For each of the seven targeted races, we combined the following facial cues from two source avatars of that race: straight eyebrows (male), rounder and lower cheeks (female), a square jaw (male), fuller lips (female), absence of lip coloring (male), and swept-back hair. Prior research shows that these blended traits collectively form a perceptually androgynous face [6]. This process was completed for two separate combinations, leading to 14 total avatars. Table 1 shows all source avatars used to create all mixed-gender avatars.

In Figure 2, headshots of all 14 avatars crafted for our study are showcased. Notably, the guidelines utilized for design and evaluation were tailored exclusively for White avatars, lacking specific considerations for Black avatars, especially in terms of hairstyles.

Table 1: A list of which source avatars from VALID [3] were used to create each mixed-gender avatar.

	Source Avatars	
	Male	Female
AIAN_A_1	AIAN_M_1	AIAN_F_1
AIAN_A_2	AIAN_M_2	AIAN_F_1
Asian_A_1	Asian_M_1	Asian_F_3
Asian_A_2	Asian_M_1	Asian_F_3
Black_A_1	Black_M_1	Black_F_3
Black_A_2	Black_M_2	Black_F_3
Hispanic_A_1	Hispanic_M_1	Hispanic_F_3
Hispanic_A_2	Hispanic_M_2	Hispanic_F_2
MENA_A_1	MENA_M_1	MENA_F_1
MENA_A_2	MENA_M_2	MENA_F_3
NHPI_A_1	NHPI_M_1	NHPI_F_3
NHPI_A_2	NHPI_M_2	NHPI_F_3
White_A_1	White_M_3	White_F_3
White_A_2	White_M_3	White_F_2

In response, we opted for a swept-back hairstyle, acknowledging the significance of hair texture for ethnic identity, as emphasized by ethnic representatives in the original VALID study [3].

4 AVATAR VALIDATION STUDY

4.1 Procedure

This procedure received approval from our university’s Institutional Review Board (IRB). The study was conducted through a single online Qualtrics survey, with an average completion time of approximately 5 minutes. Our methodology closely followed the approach utilized for the original VALID avatars [3]. Participants were initially presented with definitions of seven distinct racial identities according to the U.S. Census Bureau (AIAN, Asian, Black, Hispanic, MENA, NHPI, and White) [5] before providing their demographic details, including age, gender, and ethnicity. Subsequently, they evaluated and categorized 14 avatars with mixed-gender facial features, each depicted from three perspectives: a 45°left headshot, a direct (0°) headshot, and a 45°right headshot. Similar to VALID’s procedure, all avatars were attired in plain gray shirts and illuminated by a soft white (#FFFEF5) directional light with an intensity of 1.0, against a light gray (#7F7F7F) background. Participants viewed one avatar at a time, and the presentation order was randomized.

For each of the 14 avatars, participants were asked to categorize the perceived race and gender using two questions: a multiple-choice “select all that apply” question to identify the avatar’s race (following U.S. Census Bureau guidelines for racial surveying [5]), and a slider question gauging whether the avatar appeared more masculine, androgynous, or feminine (similar to Nag and Yalcin’s methodology [6]). The slider initially rested at the midpoint between “Male” and “Female” anchors, and participants were instructed to adjust its position to reflect their perception of the avatar’s gender. A leftward slide indicated a more masculine perception, while a rightward slide indicated a more feminine perception. Upon completion, participants were compensated with a \$1 USD gratuity through Prolific.

4.2 Participants

We recruited 68 participants (33 men, 33 women, 2 non-binary) from the online Prolific marketplace to take part in the study. Participants were aged 19-40 ($M=30.26$, $SD=9.90$) with the following racial demographics breakdown: 10 AIAN, 10 Asian, 10 Black, 10 Hispanic, 10 MENA, 10 White, and 8 NHPI. Table 2 shows the demographic breakdown of our participants.



Figure 2: Headshots of all 14 mixed-gender avatars that we created for the study.

Table 2: Demographic breakdown of our participants, showing race, gender (male, female, or nonbinary), age (mean and standard deviation), and country of residence.

Race	Gender	Age	Country
AIAN	5M, 5F	27.7 (5.8)	Austria (1), Canada (6), U.S. (2), Ireland (1)
Asian	5M, 5F	31.5 (12.1)	China (4), Canada (1), Hong Kong (1), Mongolia (1), Portugal (1), Taiwan (1), U.K. (1)
Black	5M, 5F	24.2 (3.5)	South Africa (8), Zimbabwe (1), Democratic Republic of Congo (1)
Hispanic	4M, 5F, 1NB	27.6 (8.9)	Mexico (8), Chile (1), Portugal (1)
MENA	5M, 5F	30.5 (12.1)	Turkey (4), Lebanon (3), Canada (1), Iran (1), Syrian Arab Republic (1)
NHPI	5M, 3F	34.5 (8.9)	New Zealand Maori (6), Fiji (1), Samoa (1)
White	5M, 4F, 1NB	30.4 (12.4)	Portugal (4), Canada (1), France (1), Latvia (1), Netherlands (1), Poland (1), U.K. (1)

4.3 Results

To validate the avatars’ race classifications, we adopted the methodology outlined in the initial VALID research [3]. This involved conducting a Cochran’s Q test across all participants at a 95% confidence level to detect significant differences in their responses. Upon identifying differences, pairwise comparisons between races were performed using Dunn’s test. For each avatar, a race label was assigned if it was predominantly selected by the majority of participants (i.e., chosen by over 50% of participants) and exhibited a statistically significant difference compared to other race choices, while not significantly below any other race. We also computed the agreement rate per avatar across all seven racial categories. The agreement rate represents the percentage of participants who identified the avatar as belonging to a particular race.

5 DISCUSSION

5.1 Mixed-Gender Avatars Are Racially Ambiguous

In Figure 3, we present the agreement rates for all avatars, evaluated by all 68 participants. Consistent with the VALID study [3], our Asian, Black, and White avatars received clear and distinct identifications as their source avatars. However, compared to the original study, the racial categorization of all other avatars exhib-

ited heightened ambiguity. Notably, the AIAN, Hispanic, MENA, and NHPI avatars generated significant confusion, with participants often categorizing them as a blend of all seven races (excluding Black). Intriguingly, despite the uncertainty in racial categorization, participants rarely selected “Other” for any avatar; only one instance of this choice was recorded. This suggests that while participants hesitated to assign avatars to the listed races, they also refrained from considering them as belonging to an unlisted race.

These findings suggest a potential relationship between gender perception and race perception. We posit that this connection may arise from a lack of perceptual expertise or familiarity, which are known factors influencing racial categorization. As noted by Civile et al. [2], perceptual expertise, facilitated by perceptual learning, moderates the reduced ability to recognize faces from different racial backgrounds. This effect extends to mixed-gender avatars, where the lack of exposure to avatars with blended male and female facial cues in media or real-life contexts contributes to uncertainty in categorization due to limited perceptual expertise.

5.2 Reduced Choice Diversity Among Same-Race Participants

Given Do et al.’s [3] observation that individuals of the same race exhibited higher agreement rates for the source avatars, we investigated how participants of the same race identified the mixed-gender avatars. Figure 4 exclusively presents data from participants sharing the same race as the source avatars. For instance, the agreement rates for AIAN avatars include data solely from the 10 AIAN participants. However, our findings mirrored those from the overall participant pool—avatars were consistently perceived as racially ambiguous, even by participants of the same race as the source avatars.

Interestingly, we observed that certain participants exhibited lower accuracy in categorizing mixed-gender avatars of their own race compared to participants from other racial backgrounds. For instance, as depicted in Figure 4, Hispanic_A_1 was accurately identified only 30% of the time by Hispanic participants, while general participants achieved a 46% accuracy rate. This disparity may be attributed to cultural variations. Teran et al. [10] suggest that some Latinx cultural values, such as familismo, machismo, and caballerismo, could potentially impede the acceptance of gender diversity by Hispanic fathers. Our results illustrate the potential nuanced interaction between culture and perceptual processes.

However, among participants of the same race, we observed reduced choice diversity. While the overall group of 68 participants tended to categorize AIAN, Hispanic, MENA, and NHPI avatars as

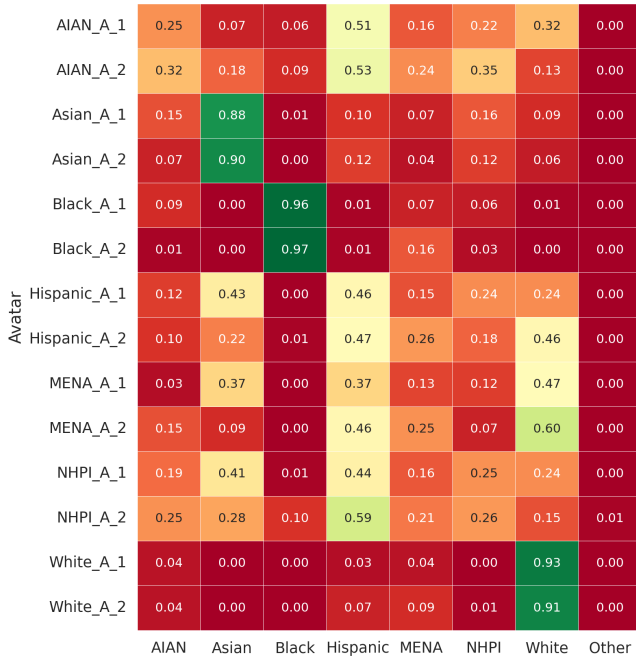


Figure 3: A heatmap of agreement rates for each avatar and its perceived race by all 68 participants.

a blend of all races (excluding Black), those of the same race tended to concentrate their categorizations on a limited set of races. These distinctions suggest that individuals of the same race may possess the ability to refine their categorizations of mixed-gender avatars to a select subset of races, unlike the broader participant pool. The reduced choice diversity among same-race participants may stem from heightened perceptual familiarity with faces of their own race. Participants demonstrated increased accuracy in narrowing down the avatar’s race to a few specific races, leveraging recognition of facial features common in their ethnic group. However, the mixture of male and female facial cues introduced a confounding factor, ultimately causing participants to perceive the avatar as racially ambiguous.

6 CONCLUSION

In conclusion, our study using the VALID Avatar Library to create mixed-gender avatars across seven racial categories uncovered intriguing insights into the nuanced relationship between gender and race perception. Despite the source avatars sharing the same racial background, the resulting mixed-gender avatars displayed heightened racial ambiguity, challenging conventional categorization. Notably, the analysis among same-race participants revealed their capacity to refine categorizations, showcasing a perceptual adaptability. These findings underscore the complex dynamics influencing the perception of mixed-gender avatars and suggest potential links between gender and race perception, rooted in limited perceptual expertise. As virtual environments become increasingly diverse, understanding these intricacies becomes imperative for creating inclusive and accurately perceived representations within digital spaces

REFERENCES

[1] S. Booth, S. Goodman, G. Kirkup, A. Gulz, and M. Haake. *Challenging Gender Stereotypes Using Virtual Pedagogical Characters*, pp. 113–132. 01 2010. doi: 10.4018/978-1-61520-813-5.ch007

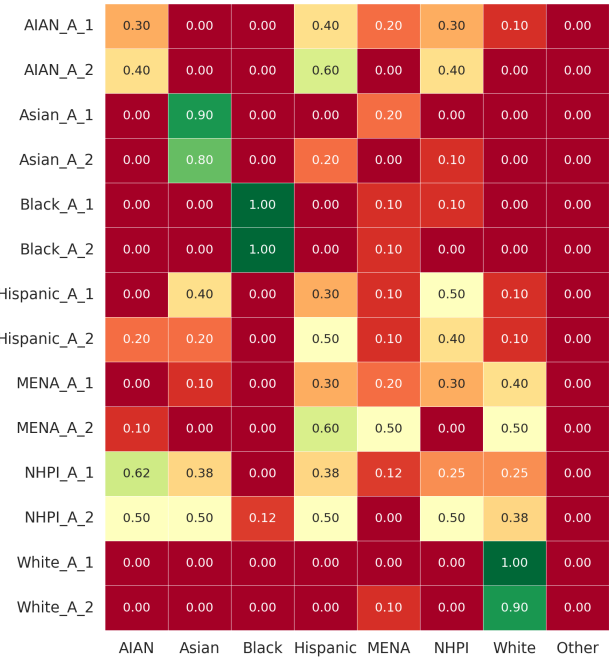


Figure 4: A heatmap of agreement rates for each avatar and its perceived race by same-race participants. For instance, the agreement rates for AIAN avatars include data solely from the 10 AIAN participants.

[2] C. Civile and I. P. L. McLaren. 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):1–10, 2022. doi: 10.1038/s41598-022-17294-w

[3] T. D. Do, S. Zelenty, M. Gonzalez-Franco, and R. P. McMahan. Valid: a perceptually validated virtual avatar library for inclusion and diversity. *Frontiers in Virtual Reality*, 4, 2023. doi: 10.3389/frvir.2023.1248915

[4] R. Ghosh, P. G. Feijóo-García, J. Stuart, C. Wrenn, and B. Lok. Evaluating face gender cues in virtual humans within and beyond the gender binary. *Frontiers in Virtual Reality*, 4:1251420, Aug. 2023. doi: 10.3389/frvir.2023.1251420

[5] K. Matthews, J. Phelan, N. A. Jones, S. Konya, R. Marks, B. M. Pratt, J. Combs, and M. Bentley. 2015 National Content Test Race and Ethnicity Analysis Report. Technical report, U.S. Department of Commerce, 2017.

[6] P. Nag and O. N. Yalçın. Gender stereotypes in virtual agents. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, IVA ’20*. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3383652.3423876

[7] A. Niculescu, D. Hofs, B. Van Dijk, and A. Nijholt. How the agent’s gender influence users’ evaluation of a QA system. In *2010 International Conference on User Science and Engineering (i-USER)*, pp. 16–20. IEEE, Shah Alam, Dec. 2010. doi: 10.1109/IUSER.2010.5716715

[8] A. Niculescu, F. Van Der Sluis, and A. Nijholt. Femininity, masculinity and androgyny: how humans perceive the gender of anthropomorphic agents. In *13th International Conference on Human-Computer Interaction, HCI International 2009*, pp. 628–632. Springer, 2009.

[9] V. Schwind and N. Henze. Gender- and Age-related Differences in Designing the Characteristics of Stereotypical Virtual Faces. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, pp. 463–475. ACM, Melbourne VIC Australia, Oct. 2018. doi: 10.1145/3242671.3242692

[10] M. Teran, R. L. Abreu, E. S. Tseung, and J. Castellanos. Latinx fathers of transgender and gender diverse people: Journey toward acceptance and role of culture. *Family Relations*, 72(4):1908–1925, 2023. doi: 10.1111/fare.12799