Collaborating with my Doppelgänger: The Effects of Self-similar Appearance and Voice of a Virtual Character during a Jigsaw Puzzle Co-solving Task

SIQI GUO, Purdue University, USA
MINSOO CHOI, Purdue University, USA
DOMINIC KAO, Purdue University, USA
CHRISTOS MOUSAS, Purdue University, USA

Fig. 1. A user interacting with the self-similar virtual character in our virtual reality application.

The research community has long been interested in human interaction with embodied virtual characters in virtual reality (VR). At the same time, interaction with self-similar virtual characters, or virtual doppelgängers, has become a prominent topic in both VR and psychology due to the intriguing psychological effects these characters can have on people. However, studies on human interaction with self-similar virtual characters are still limited. To address this research gap, we designed and conducted a 2 (appearance: self-similar vs. non-self-similar appearance) × 2 (voice: self-similar vs. non-self-similar voice) within-group study ($N = 25$) to explore how combinations of appearance and voice factors influence participants' perception of virtual characters. During the study, we asked participants to collaborate with a virtual character in solving a VR jigsaw puzzle. After each experimental condition, we had participants complete a survey about their experiences with the virtual character. Our findings showed that 1) the virtual characters' self-similarity in appearance enhanced the sense of co-presence and perceived intelligence, but it also elicited higher eeriness; 2) the self-similar

Authors' addresses: Siqi Guo, guo477@purdue.edu, Purdue University, 401 N. Grant St., West Lafayette, Indiana, USA, 47907; Minsoo Choi, choi714@purdue.edu, Purdue University, 401 N. Grant St., West Lafayette, Indiana, USA, 47907; Dominic Kao, kaod@purdue.edu, Purdue University, 401 N. Grant St., West Lafayette, Indiana, USA, 47907; Christos Mousas, cmousas@purdue.edu, Purdue University, 401 N. Grant St., West Lafayette, Indiana, USA, 47907.

ACM 2577-6193/2024/5-ART4
https://doi.org/10.1145/3651288

voices led to higher ratings on the characters' likability and believability; however, they also induced a more eerie sensation; and 3) we observed an interaction effect between appearance and voice factors for ratings on believability, where the virtual characters were considered more believable when their self-similarity in appearance matched that of their voices. This study provided valuable insights and comprehensive guidance for creating novel collaborative experiences with self-similar virtual characters in immersive environments.

CCS Concepts: • **Human-centered computing** → **Virtual reality**; **User studies**.

Additional Key Words and Phrases: Virtual Reality, Virtual Characters, Collaboration, Doppelgänger, Self-similar Appearance, Self-similar Voice, Task Co-solving, Jigsaw Puzzle

ACM Reference Format:
Siqi Guo, Minsoo Choi, Dominic Kao, and Christos Mousas. 2024. Collaborating with my Doppelgänger: The Effects of Self-similar Appearance and Voice of a Virtual Character during a Jigsaw Puzzle Co-solving Task. *Proc. ACM Comput. Graph. Interact. Tech.* 7, 1, Article 4 (May 2024), 23 pages. https://doi.org/10.1145/3651288

1 INTRODUCTION

According to the Cambridge Dictionary, *a doppelgänger*[1] (we use the term doppelgängers interchangeably with "self-similar virtual characters" in this paper) is "a spirit that looks exactly like a living person, or someone who looks exactly like someone else but who is not related to that person." These entities in interactive media, or virtual versions of the self, have attracted extensive research from various disciplines over the past few decades due to the fascinating psychological effects they induce in humans [Aymerich-Franch and Bailenson 2014; Aymerich-Franch et al. 2012; Gorisse et al. 2018, 2019; Kammler-Sücker et al. 2021; Kleinlogel et al. 2021]. Meanwhile, rapidly advancing technologies have significantly enhanced the accessibility of hardware and software for creating virtual characters with unprecedentedly high levels of realism. Researchers widely use technologies such as 3D scanning and rapid avatar creation software to generate digital doppelgängers, bringing their existence beyond the realm of science fiction [Hatada et al. 2019].

Previous research has demonstrated the influence of virtual characters' visual properties on human mental states and behaviors. A virtual character's visual representation can increase perceptions of trustworthiness and confidence in its potential to influence the real world [Kim et al. 2018; Mousas et al. 2018, 2021]. The degree of embodiment and anthropomorphism in the representation of virtual characters plays a role in this perception. For example, research in human-robot interaction has shown that humanoid robots, due to their human-like appearance, are popular and can foster emotional connections with humans [Siau and Wang 2018]. Visual representation, including the rendering style of a virtual character's full body and the appearance of virtual hands, impacts human behaviors [Cui and Mousas 2023; Nelson et al. 2022]. Studies have also established the influence of vocal properties on dynamics between humans and virtual characters, enhancing aspects like social presence and trust-building [Chérif and Lemoine 2019].

Given the two fundamental dimensions (appearance and voice) commonly identified in previous research that contribute to human perception of a virtual character, we conducted a study to explore how participants' self-similarity (in appearance and voice) with a virtual counterpart affects their interaction (see Fig. 1). While researchers have extensively explored the topic of intelligent virtual characters [Norouzi et al. 2018], studies focusing on interactions between humans and their doppelgängers remain limited. To address this research gap, we examined the effect of virtual characters' self-similarity on collaborative jigsaw puzzle solving tasks. We selected the jigsaw puzzle task for its ability to engage multiple cognitive functions, including visual perception, constructional praxis, and mental rotation [Fissler et al. 2018]. Additionally, the popularity of jigsaw puzzles as a leisure activity is likely to enhance participants' engagement within the virtual reality setting,

[1] https://dictionary.cambridge.org/us/dictionary/english/doppelganger'q=doppelg%C3%A4nger

leading to more natural interactions with the virtual characters. We conducted a within-group study with $N = 25$ participants, using a 2 (appearance: self-similar vs. non-self-similar appearance) $\times$ 2 (voice: self-similar vs. non-self-similar voice) design. In the self-similar conditions, participants interacted with a doppelgänger whose appearance, voice, or both closely resembled their own. Meanwhile, for the non-self-similar conditions, we used virtual characters with generic appearance, voice, or both for the participants to interact with.

We organized our paper in the following sections. In Section 2, we discussed the previously conducted research related to our study. In Section 3, we detailed the methodology adopted in this study. In Section 4, we demonstrated the result of this study. Then we discussed the findings of our study in Section 5. In Section 6, we discuss design recommendations for self-similar virtual characters and the limitations of this project in Section 7. Finally, we draw our conclusions and recommendations for future research in Section 8.

2 RELATED WORK

2.1 Interacting with Virtual Characters

The research community has actively investigated human interactions with virtual characters for a long time [Cui et al. 2021; Krogmeier and Mousas 2020; Krogmeier et al. 2019; Mazumdar and Mousas 2021; Nelson et al. 2023]. Among others, Blascovich et al. [Blascovich et al. 2002] demonstrated how social interactions in virtual environments can elicit responses similar to real-world interactions, Schultze [Schultze 2010] provided critical insights into the role of virtual characters as social actors, and Fox et al. [Fox et al. 2015] offered instrumental insights into the psychological effects of embodiment in virtual environments. Likewise, Kim and Biocca [Kim and Biocca 2018] focused on the sensory and cognitive aspects of interacting with virtual characters, offering a deeper understanding of immersive experiences. These works collectively highlight the significant strides made in comprehending the complex dynamics of human-virtual character interactions, laying the groundwork for current and future explorations in this rapidly evolving field.

Building on the seminal work of Nass et al. [Nass et al. 1994], who posited that interactions with virtual characters are inherently social, recent studies have continued to advance our understanding in this field. Pathi et al. [Pathi et al. 2019] revealed that people often follow the same social norms when engaging with virtual characters as they would in human interactions. Lee et al. [Lee et al. 2006] further supported this argument by demonstrating that their study participants could recognize personality traits in non-humanoid robots through verbal and non-verbal cues. Moreover, a survey by Kyrlitsias and Michael-Grigoriou [Kyrlitsias and Michael-Grigoriou 2022] discussed the various applications and advantages of immersive virtual reality (VR) technologies, emphasizing the need for people to perceive and react socially to virtual characters in a realistic manner.

Researchers have conducted a considerable number of studies on the behavioral cues of robots and virtual characters, as well as their impact on perceived intelligence. Häring et al. [Häring et al. 2012] suggested that gazing significantly improves human-robot interaction when combined with pointing gestures. Ullman et al. [Ullman et al. 2014] found that robots exhibiting complex behaviors, such as cheating, were considered more intelligent, supporting earlier findings by Short et al. [Short et al. 2010] that such behaviors can enhance user engagement. Guadagno et al. [Guadagno et al. 2007, 2011] extended these insights by suggesting that virtual characters with a high degree of behavioral realism, including the expression of non-verbal cues like smiles, have a stronger social influence and are evaluated more positively by their study participants.

Researchers have extensively studied the emotional aspect of interaction with virtual characters. Qu et al. [Qu et al. 2014] observed that the emotional states of virtual characters could trigger corresponding emotional responses in their study participants. This interaction is enriched by the virtual characters' non-verbal communication, with studies by Salem et al. [Salem et al. 2011] and Schrammel et al. [Schrammel et al. 2009] highlighting the significance of non-verbal social behaviors, such as facial expressions indicating threats, in influencing human attention and perceptions. In a study by Giuliani and Knoll [Giuliani and Knoll 2011], their study participants demonstrated considerable flexibility in their collaboration with robots in different roles. Lee et al. [Lee et al. 2015] provided further insight into how the autonomy and human-like appearance of artificial agents lead to more positive perceptions of their intelligence and trustworthiness, highlighting the importance of anthropomorphism in these interactions. The increased accessibility of VR has introduced new dimensions to human-agent interaction. Smith and Neff [Smith and Neff 2018] suggested that embodied VR enhances communication efficiency over non-embodied alternatives. This was confirmed by Fribourg et al. [Fribourg et al. 2018], who argued that embodying an avatar in VR heightens user engagement.

2.2 Appearance of Virtual Characters

The visual representation of virtual characters plays an essential role in the dynamics of human-agent interaction, significantly impacting trust, engagement, and perceived intelligence. Siau and Wang [Siau and Wang 2018] suggested that the visual representation of artificial intelligence (AI) and robots is crucial for initial trust-building with people. Weitz et al. [Weitz et al. 2019] indicated that autonomous systems are deemed more trustworthy when represented by a virtual character. Kim et al. [Kim et al. 2018] further suggested that an intelligent agent's visual embodiment bolsters their participants' engagement, social richness, presence, and confidence in the agent.

The degree of anthropomorphic realism directly relates to the level of social influence exerted by a virtual human, as observed by Jun and Bailenson [Jun and Bailenson 2020]. Volonte et al. [Volante et al. 2016] analyzed emotional responses in VR interpersonal scenarios, finding that visually realistic virtual humans elicited fewer negative effects than their less realistic counterparts. Nelson et al. [Nelson et al. 2022] highlighted that the appearance of virtual characters could sway their participants' social behavior, influencing actions such as avoidance.

The level of anthropomorphism acts as a double-edged sword. As Nowak and Biocca [Nowak and Biocca 2003] suggested, it can lead to higher expectations of the virtual character's capabilities, potentially lowering the sense of presence if those expectations are unmet. Hegel et al. [Hegel et al. 2008] discovered that their participants attribute more intelligence to anthropomorphic agents, affirming the link between appearance and perceived cognitive abilities. Interestingly, Kim et al. [Kim et al. 2020] found that collaboration with an embodied virtual character not only enhanced task performance but also reduced the perceived task load in comparison to a disembodied one. This highlights the importance of embodiment in the perceived efficiency and comfort during interactions.

The concept of the *uncanny valley*, introduced by Mori et al. [Mori 1970], stated that an individual's reaction to a robot would transition sharply from sympathy to disgust at the point where it nearly achieves human-likeness, but does not fully reach a true-to-life appearance. A study by Stein and Ohler [Stein and Ohler 2017] revealed that their study participants felt stronger eeriness when interacting with empathic characters perceived as autonomous artificial intelligence, possibly due to the perceived threat to the unique status of human emotional experiences. McDonnell et al. [McDonnell et al. 2012] addressed the visual appeal of characters, concluding that while highly realistic and cartoony render styles are equally appealing, intermediate levels of realism are seen as unattractive, which highlights the complicated nature of people's preferences in character design.

The *Proteus* effect, introduced by Yee and Bailenson [Yee and Bailenson 2007], suggests that an individual's behavior aligns with their digital self-representation. Pan and Steed [Pan and Steed 2017] found that embodying a self-avatar improves collaborative outcomes in virtual environments.

Similarly, De Rooij et al. [De Rooij et al. 2017] observed that embodying a self-similar avatar can positively moderate their participants' creativity, while Hooi and Cho [Hooi and Cho 2014] suggested that avatar self-similarity enhances self-presence and, consequently, people's self-disclosure. These studies highlight the profound influence of avatar appearance on user experience in virtual reality.

2.3 Voice of Virtual Characters

The vocal characteristics of virtual characters are critical in shaping user interaction, affecting perceptions of their understanding, expressiveness, trustworthiness, and social presence. Cabral et al. [Cabral et al. 2017] emphasized that human voices are perceived as more understandable and expressive than synthesized voices. Craig et al. [Craig et al. 2019] extended this notion by identifying the human voice as inherently more trustworthy. Moreover, Chérif and Lemoine [Chérif and Lemoine 2019] found that human voices elicit a stronger sense of social presence within virtual environments. The influence of voice extends beyond its acoustic quality, as it also carries social and gender cues. Goodman and Mayhorn [Goodman and Mayhorn 2023] suggested that the gender perceived in a virtual assistant's voice can affect its trustworthiness, indicating the influence of social categorizations on voice perception. Liew et al. [Liew et al. 2023] presented findings that voice tone, particularly enthusiasm, can significantly impact learners' emotions and cognitive load, suggesting that the voice of virtual characters could be an essential factor in educational and training settings.

The interaction between voice and appearance has also received extensive attention in recent years. Zibrek et al. [Zibrek et al. 2021] and Higgins et al. [Higgins et al. 2022] demonstrated that inconsistencies between a virtual character's realistic appearance and an unrealistic voice can elicit altered emotional responses from participants. Ferstl et al. [Ferstl et al. 2021] argued in favor of a highly realistic voice, associating it with increased likability and perceptions of human likeness. The pursuit of maximizing realism in all channels (e.g., voice, motion) is advocated by Parmar et al. [Parmar et al. 2022], who recommended that virtual character designs should aim for realism in all aspects, including voice and motion, rather than seeking consistency. Choi et al. [Choi et al. 2023] suggested that other than realism, the human likeness of voice significantly enhanced participants' perception of a non-human virtual character, as they rated the anthropomorphism higher for the same character with a human voice than with a robot voice.

2.4 Self-similarity of Virtual Characters

The degree of self-similarity in virtual characters can influence user engagement and performance. Vugt et al. [Vugt et al. 2008] discovered that participants preferred and felt more connected to an assistant agent that mirrored their facial features. Additionally, Kao et al. [Kao et al. 2021b] found that an avatar with a self-similar voice notably improved participants' task performance. However, the effects of self-similarity may vary depending on the context, as Wauck et al. [Wauck et al. 2018] observed that in their gaming scenario, self-similar avatars did not result in a marked difference in user experience. The exploration of self-similarity in virtual characters continues to evolve. Recently, Kim et al. [Kim et al. 2023] uncovered that high levels of self-similarity in avatars significantly enhance the sense of embodiment and social presence. Notably, their study highlighted the crucial role of avatar voice in enriching the social VR experience.

The sense of embodiment is also related to the customization of virtual representations. Waltemate et al. [Waltemate et al. 2018] found that personalized avatars substantially enhance the sense of body ownership, presence, and a feeling of dominance over generic avatars. Similarly, Praetorius and Görlich [Praetorius and Görlich 2020] stated that avatars reflecting participants' appearance or characteristics foster a greater personal connection and mental closeness between participants and the virtual representation. In recent research, Shin et al. [Shih et al. 2023] also identified a positive feeling induced by self-similar avatars, concluding that self-similarity in participants' visual representations enhances persuasiveness. Furthermore, Bailenson and Segovia [Bailenson and Segovia 2010] introduced the concept of a virtual doppelgänger as an autonomous virtual replica of the self, adding an intriguing dimension to the topic of self-similarity in virtual characters. Doppelgängers, acting independently of the user, raise questions about autonomy, identity, and self-perception within the discipline of virtual reality.

2.5 Research Questions

Considering the impact of appearance and voice on human perceptions of a virtual character, we aimed to dissect the nuances of how these factors influence human perception and interaction in a virtual environment. To do so, we planned to answer the following research questions:

- **RQ1:** How does collaboration with a self-similar virtual character affect participants' sense of co-presence?
- **RQ2:** How does collaboration with a self-similar virtual character influence participants' attentional allocation?
- **RQ3:** How does collaboration with a self-similar virtual character impact participants' perceived intelligence ratings?
- **RQ4:** How does collaboration with a self-similar virtual character affect participants' intelligence comparison ratings?
- **RQ5:** How does collaboration with a self-similar virtual character influence participants' eeriness ratings?
- **RQ6:** How does collaboration with a self-similar virtual character impact participants' likability ratings?
- **RQ7:** How does collaboration with a self-similar virtual character affect participants' believability ratings?
- **RQ8:** How does collaboration with a self-similar virtual character influence participants' perceived anthropomorphism ratings?

2.6 Contributions

With this study, we contribute to the research community by expanding current knowledge of how people perceive their digital doppelgängers in terms of appearance and voice match in immersive environments. We provide a deeper understanding of the effects of self-similarity in virtual characters. Our findings play a critical role in advancing the understanding of how virtual doppelgängers impact human perceptions and offer valuable insights for developing engaging virtual reality experiences.

3 MATERIALS AND METHODS

3.1 Participants

We conducted an *a priori* power analysis to determine the sample size for this study using G*Power v. 3.1 software [Faul et al. 2007]. For our 2 (appearance: self-similar vs. non-self-similar appearance) × 2 (voice: self-similar vs. non-self-similar voice) within-group study, a small effect size of $f = .25$ [Cohen 2013], and an $\alpha = .05$, to achieve an 80% power ($1-\beta$ error power), our analysis recommended a minimum of 24 participants. We recruited 25 participants through emails sent to the students' listservs at our university and in-class announcements. Among the 25 participants (age: $M = 21.04$, $SD = 3.40$), seven were female, 16 were male, one was non-binary, and one preferred not to say. Our participants volunteered to take part in this study without receiving monetary compensation.

3.2 Virtual Reality Application

We developed the virtual reality application of our study using Unity game engine v. 2020.3.20, along with the Oculus Integration Toolkit. We used Meta's Quest 2 head-mounted display, connected to a Dell Alienware Aurora R7 desktop computer (Intel Core i7, NVIDIA GeForce RTX 2080, 32GB RAM) for both the development process and the study. The application features a virtual environment resembling a living room in a semi-realistic style (see Fig. 2). The environment was lit with a directional light to resemble sunlight coming from the windows, complemented by an additional light source inside the room. Inside the living room, the participant and the virtual character were seated at a desk, with the virtual character located on the participant's right side.

(a) (b) (c)

Fig. 2. The virtual environment we used in our study from two different viewpoints (Fig. 2a and 2b), and the table with the puzzle pieces (Fig. 2c).

In the experiment, the participant found puzzle pieces placed at random positions on the table. We designed a puzzle with 25 jigsaw puzzle pieces to ensure a moderate level of difficulty and set the dimensions of the puzzle pieces at 4×4 cm. The initial position of the puzzle pieces was consistent across all conditions. Additionally, on the table, we placed a semitransparent board for the puzzle co-solving process, and next to the game board, a completed puzzle was positioned for reference.

We scripted the virtual characters to assist participants in placing the puzzle pieces in the right spot. In the VR application, we implemented a virtual character that could co-solve the jigsaw puzzle with participants, driven by a loop-based function. This function enabled the virtual character to pick up a puzzle piece and place it on the puzzle board, continuing until the virtual character and participant completed the jigsaw puzzle co-solving process. Specifically, the function determined which puzzle piece the virtual character should pick up from the group of unsolved pieces on the table. It also identified the correct spot for the picked puzzle piece and guided the virtual character to place it there.

Throughout the jigsaw puzzle co-solving process, we integrated casual dialogues (see Table 1 and the accompanying video) such as *"Let's solve this puzzle together!"* to mimic a natural conversation between two people collaborating on a task. The study participants responded to the dialogues by selecting from predesigned answers presented on a pop-up interface using the VR controllers. To achieve lifelike movements in the virtual characters, we used the Salsa LipSync Suite[2] from the Unity Asset Store for lip-sync animation. We implemented an inverse kinematics solver for the hand-reaching and upper-body movements. Moreover, to enhance the animation quality, we assigned idle motions to the virtual characters and incorporated gaze routines, ensuring their behavior appeared lifelike during the interaction.

[2]https://assetstore.unity.com/packages/tools/animation/salsa-lipsync-suite-148442

Table 1. We implemented dialogs in our application, along with responses that participants could choose from. We also indicate at which timesteps of the jigsaw puzzle co-solving process each dialog appears.

#	Dialogue Content	Options for the Participant	Timestep
1	Hi, my name is [name of the participant]. I am happy to meet you. How was your day?	1) It was good! 2) A little bit bored.	At the beginning of the puzzle co-solving process.
2	Okay, let's solve this puzzle together!	1) Good, let's do it! 2) No, I don't want to play.	At the beginning of the puzzle co-solving process.
3	This puzzle looks hard to solve.	NA	At the beginning of the puzzle co-solving process.
4	I enjoy solving puzzles.	NA	In the first quarter of the puzzle co-solving process.
5	Let me think where this puzzle piece goes.	NA	Repeated a few times during the entire puzzle co-solving process.
6	We are halfway done! Let's cheer up!	1) Nice! Let's go for it! 2) I don't want to play anymore.	When half of the puzzle is solved.
7	It looks easy now.	NA	In the third quarter of the puzzle co-solving process.
8	There are four pieces left.	NA	When four jigsaw puzzle pieces are left unsolved.
9	Good job! We did it!	1) Good job! 2) Bye.	At the end of the puzzle co-solving process.

3.2.1 Virtual Characters. For our study, we used two sets of virtual characters. We downloaded a male (Male_Adult_01) and a female (Female_Adult_01) virtual character from Microsoft's Rocketbox library [Gonzalez-Franco et al. 2020] for the non-self-similar conditions (see Fig. 3a and 3b). We chose these characters to match the gender of the participants in the experimental setup. We assigned participants who identified as non-binary a virtual character corresponding to their biological sex.

(a) (b) (c)

Fig. 3. The female (Fig. 3a) and male (Fig. 3b) virtual characters we used for the non-similar appearance conditions, and an example of a self-similar virtual character (Fig. 3c).

For the self-similar appearance conditions, we created virtual characters using the Character Creator v. 4.31[3] software, with the Headshot V2[4] plugin (see Fig. 4 for an example of a self-similar virtual character that we created). We generated the models using photos of the participants, which we took during their initial appointment following enrollment in the study. We used a Fujifilm XT-4 camera paired with a 35mm F2 lens to capture these images, ensuring a high-resolution output of 6240×4160 pixels. Furthermore, we took all photographs under identical lighting conditions and with the same background to maintain consistency. In cases where the Character Creator couldn't

[3] https://www.reallusion.com/character-creator/
[4] https://www.reallusion.com/character-creator/headshot/

generate certain accessories (e.g., a beard) from a photo, a 3D modeler with over four years of experience manually added them to ensure maximized self-similarity. Moreover, we knew that body types could affect participants' perception of the virtual characters [Lam et al. 2023]; therefore, the 3D modeler manually adjusted them in Character Creator to reflect the participants' approximated height and weight. We believe the impact of the minor mismatch between the self-similar characters and the participants' precise body types was minimal. This is because, during the interaction, both the participant and the virtual character remained seated, with the virtual table blocking the legs and most of the torso of the virtual character.

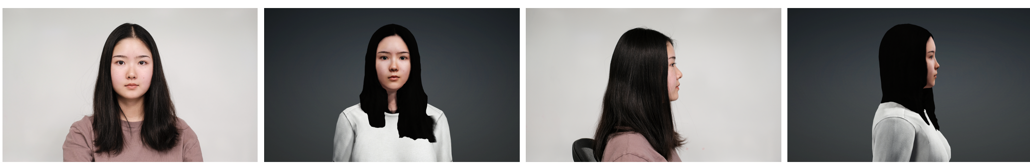

Fig. 4. Comparative front and side views of self-similar character models generated from a human's photos in Character Creator v4.31 software.

3.2.2 Voices. In our study, we used PlayHT,[5] a text-to-speech service, for the virtual characters' speech. To obtain gender-matched, non-self-similar voices, we chose predefined voice models. For non-binary participants, we selected voices matching their biological sex. Specifically, we opted for the American English-speaking voice models (Evelyn and Hudson voice models for the female and male virtual characters, respectively) from PlayHT's database. For self-similar voices, we created audio clips using voice cloning technology, based on recordings from the participants. We recorded these by providing participants with Tim Burton's short story *"The Melancholy Death of Oyster Boy"* [Burton 1997] and asking them to read it naturally and moderately. During the recording, we ensured a quiet environment in the research lab to minimize background noise. Where necessary, we used Descript[6] v. 79.1.2 for additional noise reduction. After collecting the audio clips, we employed PlayHT's voice cloning service to generate character speeches from the pre-designed dialogues (see Table 1).

3.2.3 Sanity Checks. After creating the self-similar avatars and voices, we conducted a two-phase sanity check. In the first phase, three animation experts (two men and one woman), aged 26-38 (age: $M = 31.67$, $SD = 6.03$), with over eight years of experience in 3D graphics on average (experience: $M = 8.67$, $SD = 5.69$), voted on whether the designed characters looked and sounded similar to the participants. If any expert disagreed, we requested the 3D modeler to revise the virtual character based on their suggestions. We followed a similar process for the synthesized voices, repeating it until all three experts were satisfied with the edits.

In the second phase, we evaluated the appearances and voices of the self-similar virtual characters with the participants to ensure that the avatar and voice matched their perceptions. After creating the self-similar virtual characters, we distributed a short survey to each participant with their own self-similar avatar. Participants observed their virtual doppelgänger for a few seconds and then anonymously rated three statements on a 7-point Likert scale: 1) Q1: The virtual character looked like me. (1=Not at all; 7=Totally); 2) Q2: The virtual character sounded like me. (1=Not at all; 7=Totally); and 3) Q3: The appearance and voice of the virtual character resemble me. (1=Not at all; 7=Totally). The survey results showed that the virtual characters' self-similarity was convincing,

[5]https://play.ht/
[6]https://www.descript.com/

with relatively high mean values, while all three statements exceeded the midpoint of the scale: Q1 (M = 5.38, SD = 1.03), Q2 (M = 5.25, SD = 1.07), and Q3 (M = 4.75, SD = 1.81). Cronbach's alpha indicated acceptable reliability, $\alpha = .72$. Based on these results, we concluded that the self-similar virtual characters were well-designed.

3.3 Experiment Conditions

We developed four experimental conditions to investigate how the virtual character's self-similarity in voice and appearance affected participants' perception of them and whether an interaction existed between the two factors (self-similarity of appearance and voice). Specifically, we examined the following four conditions:

- **Non-self-similar appearance and non-self-similar voice (NANV):** In this experimental condition, we used the gender-matched character from Microsoft's Rocketbox library for the virtual character's visual representation and generated the speech using the gender-matched voice model on PlayHT.
- **Non-self-similar appearance and self-similar voice (NASV):** In this experimental condition, we used the gender-matched character from Microsoft's Rocketbox library for the virtual character's visual representation and generated the speech through voice cloning using the participant's audio.
- **Self-similar appearance and non-self-similar voice (SANV):** In this condition, we generated the virtual character's visual representation using the participant's photo and generated the speech using the gender-matched voice model on PlayHT.
- **Self-similar appearance and self-similar voice (SASV):** In this condition, we generated the virtual character's visual representation using the participant's photo and generated the speech through voice cloning using the participant's audio.

3.4 Survey

We created a survey to explore the effects of interacting with self-similar virtual characters, collecting subjective self-reported ratings from participants. This approach helped us understand their perceptions of the characters under various conditions. The survey, detailed in Table A1 in Appendix A, comprises 42 items across eight variables. We included six items to measure co-presence and six for attentional allocation, both sets developed by Biocca et al. [Biocca et al. 2001]. Six items assessed perceived intelligence, and another six gauged perceived anthropomorphism, both sets formulated by Moussawi and Koufaris [Moussawi and Koufaris 2019]. The survey also contained an item for intelligence comparison, created by us, along with three items to measure perceived eeriness by Zibrek et al. [Zibrek et al. 2018], 11 items from Reysen's likability scale [Reysen 2005], and three items evaluating the virtual character's believability from Lam et al. [Lam et al. 2023]. Additionally, we incorporated an open-ended question for participant feedback. We would like to note that we altered some of the items of our survey to match the scope of our study. We distributed this survey using Qualtrics, an online survey platform.

3.5 Procedure

Volunteer participants scheduled two separate appointments for this study. During the first appointments, we greeted the participants in the lab room with an oral introduction to the study and instructed them to complete a pre-screening form to determine their eligibility. We provided a consent form, approved by our university's Institutional Review Board (IRB), to participants who reported no symptoms, such as a history of severe motion sickness. We proceeded to the next step after the participant signed the consent form.

In the subsequent step, the researcher took headshots of the participants. We ensured consistency by taking photos at the same angle, distance, location, and under identical lighting conditions. After the photography session, the researcher directed the participants to read several paragraphs from a short story and record their audio for voice cloning. Once we obtained the photo and audio clip, we generated the self-similar avatar and cloned the voice, following the steps outlined in Sections 3.2.1 and 3.2.2. We then imported the avatar and audio clips into the VR application for self-similar voice and appearance conditions, as depicted in Fig. 3c.

During the second appointment, we asked the participants to complete a short demographics survey. We then informed them that they could take breaks, stop, and leave the study at any time without any consequences. Next, we instructed them to wear the Meta Quest 2 head-mounted display and begin the four experimental conditions. We pre-determined the sequence of these conditions using the Latin squares method [Williams 1949] to eliminate carry-over (residual) effects across the examined conditions. Before the experimental conditions, we provided our participants with a tutorial scene to introduce the controllers and ensure they understood how the jigsaw puzzle interaction worked. A previous study has shown that tutorials about virtual reality controllers could improve participants' user experience [Kao et al. 2021a]. After each condition, we requested our participants to complete the survey we developed. Upon finishing all four conditions and surveys, we expressed our gratitude to the participants, and they left the lab.

4 RESULTS

4.1 Self-reported Data

In our statistical analysis, we used the appearance and voice factors as independent variables and the self-reported ratings as dependent variables. The Q-Q plots of the residuals and the Shapiro-Wilk test at the 5% level confirmed the normality of the collected data. Thus, we performed a two-way repeated measures analysis of variance (RM-ANOVA) for each variable using IBM's SPSS software v. 25 (see Table 2 for a detailed breakdown of the results.)

Co-Presence. Our simple main effect analysis on the appearance factor (Wilk's $\Lambda = .836$, $F[1, 24] = 4.695$, $p = .040$, $\eta_p^2 = .164$) showed that participants rated their co-presence higher when collaborating with virtual characters with self-similar appearance ($M = 6.04$, $SE = .10$) than with non-self-similar ones ($M = 5.84$, $SE = .10$). However, we did not find a statistically significant main effect for the voice factor (Wilk's $\Lambda = .917$, $F[1, 24] = 2.181$, $p = .153$, $\eta_p^2 = .083$) or for the appearance × voice interaction effect (Wilk's $\Lambda = .964$, $F[1, 24] = .895$, $p = .354$, $\eta_p^2 = .036$).

Attentional Allocation. We found no statistically significant main effect for the appearance factor (Wilk's $\Lambda = .935$, $F[1, 24] = 1.655$, $p = .211$, $\eta_p^2 = .065$), the voice factor (Wilk's $\Lambda = .951$, $F[1, 24] = 1.227$, $p = .279$, $\eta_p^2 = .049$), or the appearance × voice interaction effect (Wilk's $\Lambda = .997$, $F[1, 24] = .069$, $p = .795$, $\eta_p^2 = .003$).

Perceived Intelligence. Our simple main effect analysis on the appearance factor (Wilk's $\Lambda = .800$, $F[1, 24] = 5.986$, $p = .022$, $\eta_p^2 = .200$) showed that participants rated the perceived intelligence higher for virtual characters with self-similar appearance ($M = 5.82$, $SE = .08$) compared to those with non-self-similar appearance ($M = 5.67$, $SE = .09$). However, we did not find a statistically significant main effect for the voice factor (Wilk's $\Lambda = .992$, $F[1, 24] = .182$, $p = .673$, $\eta_p^2 = .008$) or for an appearance × voice interaction effect (Wilk's $\Lambda = .977$, $F[1, 24] = .572$, $p = .457$, $\eta_p^2 = .023$).

Intelligence Comparison. We found no statistically significant main effect for the appearance factor (Wilk's $\Lambda = .999$, $F[1, 24] = .014$, $p = .908$, $\eta_p^2 = .001$), the voice factor (Wilk's $\Lambda = .971$,

$F[1, 24] = .723$, $p = .403$, $\eta_p^2 = .029$), or the appearance × voice interaction effect (Wilk's $\Lambda = .929$, $F[1, 24] = 1.839$, $p = .188$, $\eta_p^2 = .071$).

Eerie. Our simple main effect analysis on the appearance factor (Wilk's $\Lambda = .382$, $F[1, 24] = 38.860$, $p = .000$, $\eta_p^2 = .618$) revealed that participants rated the eeriness higher for virtual characters with self-similar appearance ($M = 4.67$, $SE = .13$) compared to those with non-self-similar appearance ($M = 3.84$, $SE = .12$). The main effect analysis on the voice factor (Wilk's $\Lambda = .473$, $F[1, 24] = 26.733$, $p = .000$, $\eta_p^2 = .527$) also revealed that participants rated the eeriness higher for virtual characters with self-similar voices ($M = 4.51$, $SE = .13$) compared to those with non-self-similar voices ($M = 4.01$, $SE = .14$). However, the analysis did not reveal an appearance × voice interaction effect (Wilk's $\Lambda = 1.000$, $F[1, 24] = .003$, $p = .960$, $\eta_p^2 = .000$).

Likability. We found no statistically significant main effect for the appearance factor (Wilk's $\Lambda = .882$, $F[1, 24] = 3.212$, $p = .086$, $\eta_p^2 = .118$). However, our simple main effect analysis on the voice factor (Wilk's $\Lambda = .799$, $F[1, 24] = 6.043$, $p = .022$, $\eta_p^2 = .201$) showed that participants rated the likability of virtual characters higher when exposed to self-similar voices ($M = 4.27$, $SE = .13$) compared to non-self-similar voices ($M = 3.92$, $SE = .15$). Additionally, the analysis of the appearance × voice interaction effect showed no statistically significant results (Wilk's $\Lambda = .964$, $F[1, 24] = .902$, $p = .352$, $\eta_p^2 = .036$).

Believability. We did not find a statistically significant main effect for the appearance factor (Wilk's $\Lambda = .993$, $F[1, 24] = .173$, $p = .681$, $\eta_p^2 = .007$). However, our simple main effect analysis for the voice factor (Wilk's $\Lambda = .650$, $F[1, 24] = 12.912$, $p = .001$, $\eta_p^2 = .350$) showed that participants rated the believability higher for virtual characters with self-similar voices ($M = 4.47$, $SE = .18$) compared to those with non-self-similar voices ($M = 3.49$, $SE = .20$). Additionally, we found a statistically significant appearance × voice interaction effect (Wilk's $\Lambda = .443$, $F[1, 24] = 30.176$, $p = .000$, $\eta_p^2 = .557$), indicating participants rated virtual characters as more believable when the self-similarity in voice matched that of appearance.

Perceived Anthropomorphism. We found no statistically significant main effect for the appearance factor (Wilk's $\Lambda = 1.000$, $F[1, 24] = .002$, $p = .961$, $\eta_p^2 = .000$), the voice factor (Wilk's $\Lambda = .899$, $F[1, 24] = 2.702$, $p = .113$, $\eta_p^2 = .101$), or the appearance × voice interaction effect (Wilk's $\Lambda = .921$, $F[1, 24] = 2.062$, $p = .164$, $\eta_p^2 = .079$).

4.2 Qualitative Data

We examined participants' qualitative survey responses to understand their experiences and reported interactions in our virtual reality study. We divided these responses into categories: appearance, voice, the combination of appearance and voice, and overall experience, summarizing them in the following paragraphs.

Regarding virtual characters' **appearance**, the vast majority of participants felt the self-similar virtual characters resembled their appearance. Participants P3, P5, P11, and P16 mentioned they were "seeing themselves" in self-similar appearance conditions. For instance, P3 found it *"...eerie to see myself...,"* while P11 and P16 observed that the virtual character "looked like them." Additionally, P25 noted that the self-similarity in the virtual character's appearance increased their expectation for a "more realistic lip-sync" animation.

Concerning the virtual characters' **voice**, P13 observed, *"The voice contributed to the virtual characters' believability more than the appearance."* P22 found it *"...interesting to hear the virtual character talk with a self-similar voice."* Moreover, P25 remarked, *"...the self-similar voice was mostly accurate."*

Table 2. Detailed results of our study (significant results are bold).

	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
NANV	5.82	.80	4.16	.60	5.68	.72	2.88	1.67	3.59	.88	3.81	.95	4.11	1.30	4.75	.97
NASV	5.82	.68	4.16	.68	5.65	.58	2.76	1.39	4.09	.77	4.06	.88	3.77	1.26	4.85	.83
SANV	5.93	.72	4.32	.74	5.78	.54	2.64	1.32	4.43	.94	4.02	1.22	2.87	1.29	4.63	1.07
SASV	6.15	.62	4.32	.68	5.86	.58	3.04	1.65	4.92	.87	4.48	.89	5.16	089	4.99	.94
Main Effect Appearance																
F	**4.695**		1.655		**5.986**		.014		**38.860**		3.212		.173		.002	
p	**.040**		.211		**.022**		.908		**<.001**		.086		.681		.961	
η_p^2	**.164**		.065		**.200**		.001		**.618**		.118		.007		.000	
Main Effect Voice																
F	2.181		1.227		.182		.723		**26.733**		**6.043**		**12.912**		2.702	
p	.153		.279		.673		.403		**<.001**		**.022**		**.001**		.113	
η_p^2	.083		.049		.008		.029		**.527**		**.201**		**.350**		.101	
Interaction Effect																
F	.895		.069		.572		1.839		.003		.902		**30.176**		2.062	
p	.354		.795		.457		.188		.960		.352		**< .001**		.164	
η_p^2	.036		.003		.023		.071		.000		.036		**.557**		.079	

Appearance df=1, Voice df=1, Interaction df=1, Error df=24

Notes:
(1) Co-Presence (2) Attentional Allocation (3) Perceived Intelligence (4) Intelligence Comparison
(5) Eerie (6) Likability (7) Believability (8) Perceived Anthropomorphism

Some participants also gave feedback on the **combination of appearance and voice**. P6 expressed that *"...the uncanniness varies across the conditions,"* finding the character with self-similar appearance and voice fun and strange. However, a non-self-similar appearance with a self-similar voice made P6 report a negative feeling *"...as I was hearing my voice from that felt a little violating."* Moreover, P18 noted that virtual characters with mismatched voices and appearance were less favorable *"I did not like seeing my avatar and hearing a different voice or hearing my voice with a different face."*

Lastly, participants commented on the **overall experience**. Many found it enjoyable; P4 mentioned *"...the shared manipulation of the puzzle pieces with the virtual characters was fun."* Also, P9, P14, and P23 described the experiment as "fun" and "interesting." The interaction was perceived as sophisticated and indicative of the virtual characters' intelligence. P7 called the experiment "informative," and P9 experienced an *"...illusion that the virtual character was able to adapt to their decisions."* Meanwhile, collaboration with the doppelgängers evoked an uncanny sensation for some. Participants P1, P3, P5, P6, P16, P18, and P22 described the experience as somewhat "eerie" or "uncanny."

5 DISCUSSION

The findings of this study provided valuable insights into the impact of self-similar appearance and voice of virtual characters on human interaction in an immersive virtual environment. The overarching goal was to deepen our understanding of the nuanced relationship between humans and self-similar virtual characters.

In our study, we contributed new perspectives to virtual reality research, especially regarding **co-presence** (**RQ1**). Specifically, we found that the appearance of virtual characters significantly influenced co-presence, aligning with prior research that emphasized the importance of visual similarity in enhancing virtual experiences [Biocca et al. 2003; Lee and Nass 2005]. A notable aspect of our findings was the enhanced co-presence experienced by participants interacting with their digital doppelgängers. This enhancement aligned with the theory that familiarity and self-recognition intensify social presence. This theory is supported by research indicating that humans respond positively to entities resembling themselves, known as the "mere-exposure effect" [Montoya et al. 2017; Zajonc 2001]. In virtual environments, encountering a self-similar virtual character activated a sense of familiarity and personal relevance, thereby enhancing the user's engagement and sense of presence [Bailenson et al. 2005]. Additionally, we think the psychological phenomenon of self-recognition, which activates specific brain areas associated with self-processing [Keenan et al. 2001], played a crucial role. When participants in a virtual environment interact with characters that resemble their appearance, it might trigger these self-processing mechanisms, intensifying the feeling of co-presence. Contrary to previous studies that highlighted the importance of voice in social presence [Nass and Lee 2000], our findings did not show a significant impact of voice on co-presence. These results extended the scientific discourse, highlighting the nuanced role of self-similarity in virtual character interactions in a virtual environment.

Regarding **attentional allocation** (**RQ2**), our findings were not statistically significant, contradicting a prior study where eye-tracking data showed that visual fidelity could impact participants' visual attention [Volonte et al. 2019]. This divergence may result from the distinct methodologies employed—our study used questionnaires to assess attentional allocation, whereas the referenced study utilized eye tracking for direct measurement. The cognitively demanding nature of the co-solving puzzle experience in our experiment [Nowak and Biocca 2003] might have further contributed to this discrepancy. Participants mainly focused on solving the jigsaw puzzle, leading to a uniform pattern of attentional allocation across different conditions.

In terms of **perceived intelligence** (**RQ3**), our analysis showed a significant finding for the appearance factor. Our simple main effect analysis revealed that participants rated the perceived intelligence higher for virtual characters with self-similar appearance compared to those with non-self-similar appearance. This suggests that the appearance of virtual characters can indeed affect their perceived intelligence, a notion supported by previous studies [Choi et al. 2023]. In our experiments, despite the characters consistently placing puzzle pieces correctly in all experimental conditions and demonstrating high problem-solving ability, the ratings in perceived intelligence varied based on the characters' self-similarity in appearance. The perceived intelligence ratings in the four conditions averaged 5.71 (NANV: $M = 5.68$, $SD = .72$; NASV: $M = 5.65$, $SD = .58$; SANV: $M = 5.78$, $SD = .54$; SASV: $M = 5.86$, $SD = .58$), with a higher average for conditions involving self-similar appearance. This finding extends previous studies [Choi et al. 2023; Hegel et al. 2008], as our study indicates that self-similarity in appearance plays a more critical role in this perception than previously thought.

In the investigation of **intelligence comparison** (**RQ4**), our statistical analysis revealed no significant differences. However, the self-similar appearance and voice condition (SASV) yielded slightly higher participant ratings, averaging $M = 3.04$ ($SD = 1.65$), compared to other conditions (NANV: $M = 2.88$, $SD = 1.67$; NASV: $M = 2.76$, $SD = 1.39$; SANV: $M = 2.64$, $SD = 1.32$). Despite this increase, the average scores across all conditions were below the midpoint of the scale (< 3.5), indicating that participants consistently perceived the virtual characters as less intelligent compared to themselves. This might be due to the multifaceted nature of perceived intelligence in virtual characters and suggests that high capability alone is not sufficient for an impression of high intelligence.

Additionally, this perception could be influenced by inherent human cognitive biases. The discrepancy noted in previous research between self-estimated intelligence and tested intelligence

[Holling and Preckel 2005] provides context for our results. This notion aligns with Buunk and Van Yperen [Buunk and Van Yperen 1991], suggesting people tend to have overly positive views of their abilities, like intelligence [Yamada et al. 2013], known as *Illusory Superiority*. This bias could lead participants to perceive virtual characters' intelligence as lower compared to themselves, irrespective of actual capabilities. Furthermore, *egocentrism* theory, as illustrated by Kruger [Kruger 1999], indicates a tendency to overestimate ability in easy tasks and underestimate it in difficult ones, which could have influenced participants' perceptions of virtual characters' intelligence in our study. The complexity introduced by self-similarity could have nuanced these perceptions, but the precise nature and extent of this influence remain ambiguous and warrant further exploration.

Our findings on **eeriness** (**RQ5**) suggested that both self-similar appearance and voice induced a higher sensation of uncanniness, aligning with studies on virtual doppelgängers. This eerie sensation was caused by encountering their "virtual self" with no control over behaviors [Bailenson and Segovia 2010], interpreted by participants as a threat to human emotional experience uniqueness [Stein and Ohler 2017]. Furthermore, as indicated by a previous study, high facial similarity is associated with negative responses, especially when the virtual character is perceived as unhelpful [Vugt et al. 2008]. Our results contributed to this knowledge by incorporating the voice factor, concluding that voice similarity also amplifies eeriness.

The statistical analysis shed light on the fact that participants rated **likability** higher in the presence of self-similar voices (**RQ6**) compared to non-self-similar voices. This aligns with previous research by Kao et al. [Kao et al. 2021b], which highlighted the effectiveness and plausibility of self-similar avatar voices, particularly in educational game contexts. Our study extends these observations, suggesting the broader applicability of this phenomenon: virtual characters are perceived more favorably when their vocal characteristics mirror those of the participants. Moreover, our results contribute to the ongoing discourse on avatar customization, as explored by Kao et al. [Kao et al. 2022]. Researchers in this study posited that incorporating audio customization in avatars, alongside visual customization, could enhance the user experience, but also noted a comparatively weaker effect relative to visual customization. Interestingly, our findings offer a nuanced perspective on this matter. We observed that in the context of self-similarity, vocal resemblance plays a more important role than visual similarity in boosting likability. This could be attributed to the differences in the degrees of uncanniness associated with audio and visual resemblance, as elaborated earlier in this discussion. While both self-similar visual and audio features can lead to a sense of eeriness [Männistö-Funk and Sihvonen 2018; Mori 1970], we think that the uncanny effects are less pronounced with audio, resulting in a more favorable perception of self-similar voices compared to appearances.

The results of our statistical analysis suggest that self-similar voices lead to higher ratings for the virtual characters' **believability** (**RQ7**). Additionally, the significant appearance × voice interaction effect implies that matched self-similarity in appearance and voice induces greater believability. While partially aligning with the findings of Kao et al. [Kao et al. 2022], which emphasized the integral nature of audio-visual interaction in customizable self-avatars, our study diverges by confirming the independent effect of voice. Our findings address the necessity of design consideration for the virtual characters' voices, even in the absence of visual cues. Furthermore, our results extend the research conducted by Lam et al. [Lam et al. 2023], which concluded that audio-visual congruence substantially boosts the believability of virtual characters. Considering that believability is highly correlated with audio-visual correspondence [Lam et al. 2023], our study not only validates the importance of audio-visual correspondence but also enriches the understanding of how self-similarity contributes to the believability of virtual characters.

Regarding **anthropomorphism** (**RQ8**), our findings reveal no statistically significant results, contrasting with previous studies suggesting that the appearance of virtual characters can affect their perceived anthropomorphism [Choi et al. 2023]. This divergence might be attributed to our study's experimental design. We intentionally designed virtual characters to be comparable in terms of realism, human-like attributes, and problem-solving capabilities across all experimental conditions. According to Rujiten et al. [Ruijten et al. 2019], virtual characters with such attributes occupied a midpoint on the anthropomorphism scale, indicating that participants should not have encountered excessive difficulty in attributing anthropomorphic qualities to the virtual characters in our study. However, our data analysis reveals that perceived anthropomorphism ratings in all four experimental conditions fall on the upper end of the scale (NANV: M = 4.75, SD = .97; NASV: M = 4.85, SD = .83; SANV: M = 4.63, SD = 1.07; SASV: M = 4.99, SD = .94). These consistently moderate ratings on the anthropomorphism scale across all conditions suggest that the uniform human likeness of our virtual characters in the experiment contributes to a relatively stable perception of anthropomorphism among participants. Despite the absence of statistically significant differences, this highlights the effectiveness of our design in ensuring that virtual characters are consistently perceived as somewhat anthropomorphic, regardless of the specific condition.

In summary, our results indicate that virtual characters' self-similarity in appearance boosts co-presence and perceived inteligence, while self-similar voices enhance likability and believability. Regarding believability, we observed an interaction effect where matching self-similarity in appearance and voice leads to more believable virtual characters. Meanwhile, self-similar appearance and voice induce a higher sense of eeriness compared to generic ones.

6 DESIGN RECOMMENDATIONS

In developing self-similar virtual characters, careful consideration of design elements is crucial for creating engaging and effective virtual experiences. Thus, in this section, we discuss various design recommendations to guide researchers and developers in this field.

Firstly, maintaining the consistency between the appearance and voice of the virtual characters is advisable, as this alignment significantly enhances their believability. However, the high self-similarity in the virtual characters should be implemented with caution, since our findings indicated that while self-similarity enhanced some of the examined perceptions, it elicited a heightened sense of eeriness. Therefore, researchers should carefully evaluate the target users' acceptance and preference for this duality to maximize the effectiveness and appeal of the self-similar virtual characters.

Second, in the design of self-similar virtual characters, ensuring that representation fidelity matches audio fidelity is crucial. The visual accuracy and detail in the character's appearance should be complemented by a voice that is equally realistic and nuanced. Any mismatch between a highly realistic character model and a less convincing voice, or vice versa, can disrupt the participant's believability of the virtual character. Therefore, careful attention to syncing detailed facial features and expressions with accurate voice modulation is key to creating a cohesive and engaging virtual experience. In our study, we chose to use synthesized voices to match the fidelity of the synthesized appearance of the virtual characters, emphasizing the importance of alignment between audio and visual elements.

Third, offering customization options where users can adjust the degree of self-similarity (in terms of appearance and voice) can also significantly improve user engagement and comfort. This empowerment allows users to personalize their experience to their liking, making the virtual characters more relatable and appealing. Customization can range from basic adjustments, like changing hair color or voice tone, to more advanced features, such as altering facial features or selecting from various voice modulation options.

When implemented thoughtfully, these recommendations can substantially elevate the effectiveness and appeal of self-similar virtual characters in various interactive environments. By focusing on nuanced details, developers can create virtual characters that resonate more deeply with users, fostering a sense of connection and engagement that goes beyond superficial interaction.

7 LIMITATIONS

In presenting this study's findings, it's important to note the limitations we encountered during our research. These limitations do not invalidate the results but rather provide context for their interpretation and suggestions for future research.

First, despite efforts to recruit participants from diverse cultural backgrounds and create self-similar voices for each individual, we faced challenges with the voice cloning platform, especially with foreign accents. The technology occasionally failed to replicate accents with maximized fidelity. Participant feedback indicated that the voice cloning quality was acceptable despite these inaccuracies, pointing to a critical area for future exploration.

Second, we recognize a potential limitation due to the mismatch between the participant's actual age and the perceived age of the virtual characters in the non-self-similar conditions [Richards et al. 2020]. To address this issue and minimize the potential bias, we used identical, gender-matched virtual characters across the non-self-similar conditions. We assume this consistency will substantially reduce the impact of this limitation. Additionally, we acknowledge a limitation in representing non-binary participants due to the limited availability of non-binary avatars in standard libraries, restricting our ability to provide virtual characters with which participants might more closely identify. This limitation highlights a significant gap in technological infrastructure and needs future improvements [Spiel 2021].

Third, we recognize the potential for bias induced by the disparity in visual fidelity between the virtual characters sourced from Microsoft's Rocketbox library for non-similar conditions and those created via Character Creator for self-similar conditions. Although we regard this distinction in realism as relatively subtle, it might have influenced participants' perceptions, such as assessing perceived intelligence. Future studies should consider ensuring uniformity in the realism of virtual characters, regardless of their similarity to the participant, to eliminate this confounding factor.

Fourth, another limitation related to the interaction modality. The study relied on a graphical user interface (GUI) for interactions, which may have constrained the naturalness of the interaction. Future studies might consider incorporating free-form conversation to foster a more organic and immersive user experience, potentially enhancing user engagement with virtual characters.

Last, the animation process in this study was largely automated, using tools like Unity's Salsa LipSync Suite for lip-sync animations and inverse kinematics for the hand-reaching and some upper-body movements. While this approach provided efficiency and consistency, it inherently limited the range and subtlety of the animations. Although adequate for our experimental design, the lip-sync and idle motions could be further refined. Incorporating nuanced secondary motions and more detailed expressions could significantly enhance the realism and believability of the virtual characters, especially in scenarios where more dynamic and expressive movements are crucial.

8 CONCLUSIONS AND FUTURE WORK

Our study investigated the effect of collaborating with one's virtual doppelgänger on a cognitively demanding task in a virtual environment on human perceptions. In the study, participants interacted with virtual characters that either highly resembled or differed from their appearance and voices across four experimental conditions. We then assessed our participants' perception of the virtual characters through their provided self-reported ratings. Our statistical analyses revealed key findings. It showed that the virtual characters' self-similarity in appearance enhanced the sense of co-presence and perceived intelligence, but it also elicited higher eeriness. Moreover, the self-similar voices led to higher ratings on the characters' likability and believability; however, they also induced a more eerie sensation. Lastly, we observed an interaction effect where matching self-similarity in appearance and voice leads to more believable virtual characters.

In future studies, we will focus on crafting digital doppelgängers with even higher visual and auditory fidelity, grounded in the hypothesis that heightened representational fidelity may trigger stronger psychological responses. Additionally, we plan to employ advanced animation techniques to simulate natural human movements more accurately. This advancement, along with more complex and cognitively challenging collaborative tasks, aims to uncover deeper insights into human-virtual character interactions and their emotional and cognitive impacts. Further explorations will also involve integrating language model-based free-form conversation capabilities and personalized personality traits into the virtual characters. This approach will enable these virtual characters to look and sound like their human counterparts and converse in a manner that reflects their personality traits and linguistic style. We anticipate these advancements to significantly enrich the depth of the virtual experience, thus providing a more comprehensive understanding of the nuances of human-agent interactions.

REFERENCES

Laura Aymerich-Franch and Jeremy Bailenson. 2014. The use of doppelgangers in virtual reality to treat public speaking anxiety: a gender comparison. In *Proceedings of the International Society for Presence Research Annual Conference*. Citeseer, 173–186.

Laura Aymerich-Franch, Cody Karutz, Jeremy N Bailenson, et al. 2012. Effects of facial and voice similarity on presence in a public speaking virtual environment. In *Proceedings of the International Society for Presence Research Annual Conference*. 24–26.

Jeremy N Bailenson, Andrew C Beall, Jack Loomis, Jim Blascovich, and Matthew Turk. 2005. Transformed social interaction, augmented gaze, and social influence in immersive virtual environments. *Human communication research* 31, 4 (2005), 511–537.

Jeremy N Bailenson and Kathryn Y Segovia. 2010. Virtual doppelgangers: Psychological effects of avatars who ignore their owners. *Online worlds: Convergence of the real and the virtual* (2010), 175–186.

Frank Biocca, Chad Harms, and Judee K Burgoon. 2003. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators & virtual environments* 12, 5 (2003), 456–480.

Frank Biocca, Chad Harms, and Jenn Gregg. 2001. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. In *4th annual international workshop on presence, Philadelphia, PA*. 1–9.

Jim Blascovich, Jack Loomis, Andrew C Beall, Kimberly R Swinth, Crystal L Hoyt, and Jeremy N Bailenson. 2002. Immersive virtual environment technology as a methodological tool for social psychology. *Psychological inquiry* (2002), 103–124.

Tim Burton. 1997. *The melancholy death of Oyster boy & other stories*. Harper Collins.

Bram P Buunk and Nico W Van Yperen. 1991. Referential comparisons, relational comparisons, and exchange orientation: Their relation to marital satisfaction. *Personality and Social Psychology Bulletin* 17, 6 (1991), 709–717.

Joao Paulo Cabral, Benjamin R Cowan, Katja Zibrek, and Rachel McDonnell. 2017. The Influence of Synthetic Voice on the Evaluation of a Virtual Character.. In *Interspeech*. Stockholm, 229–233.

Emna Chérif and Jean-François Lemoine. 2019. Anthropomorphic virtual assistants and the reactions of Internet users: An experiment on the assistant's voice. *Recherche et Applications en Marketing (English Edition)* 34, 1 (2019), 28–47.

Minsoo Choi, Alexandros Koilias, Matias Volonte, Dominic Kao, and Christos Mousas. 2023. Exploring the Appearance and Voice Mismatch of Virtual Characters. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 555–560.

Jacob Cohen. 2013. *Statistical power analysis for the behavioral sciences*. Academic press.

Scotty D Craig, Erin K Chiou, and Noah L Schroeder. 2019. The impact of virtual human voice on learner trust. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 63. SAGE Publications Sage CA: Los Angeles, CA, 2272–2276.

Dixuan Cui, Dominic Kao, and Christos Mousas. 2021. Toward understanding embodied human-virtual character interaction through virtual and tactile hugging. *Computer Animation and Virtual Worlds* 32, 3-4 (2021), e2009.

Dixuan Cui and Christos Mousas. 2023. Exploring the effects of virtual hand appearance on midair typing efficiency. *Computer Animation and Virtual Worlds* (2023), e2189.

Alwin De Rooij, Sarah Van Der Land, and Shelly Van Erp. 2017. The creative Proteus Effect: How self-similarity, embodiment, and priming of creative stereotypes with avatars influences creative ideation. In *Proceedings of the 2017 ACM SIGCHI*

Conference on Creativity and Cognition. 232–236.

Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods* 39, 2 (2007), 175–191.

Ylva Ferstl, Sean Thomas, Cédric Guiard, Cathy Ennis, and Rachel McDonnell. 2021. Human or Robot? Investigating voice, appearance and gesture motion realism of conversational social agents. In *Proceedings of the 21st ACM international conference on intelligent virtual agents*. 76–83.

Patrick Fissler, Olivia Caroline Küster, Daria Laptinskaya, Laura Sophia Loy, Christine AF Von Arnim, and Iris-Tatjana Kolassa. 2018. Jigsaw puzzling taps multiple cognitive abilities and is a potential protective factor for cognitive aging. *Frontiers in aging neuroscience* 10 (2018), 408085.

Jesse Fox, Sun Joo 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.

Rebecca Fribourg, Ferran Argelaguet, Ludovic Hoyet, and Anatole Lécuyer. 2018. Studying the sense of embodiment in VR shared experiences. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 273–280.

Manuel Giuliani and Alois Knoll. 2011. Evaluating supportive and instructive robot roles in human-robot interaction. In *Social Robotics: Third International Conference, ICSR 2011, Amsterdam, The Netherlands, November 24-25, 2011. Proceedings 3*. Springer, 193–203.

Mar Gonzalez-Franco, Eyal Ofek, Ye Pan, Angus Antley, Anthony Steed, Bernhard Spanlang, Antonella Maselli, Domna Banakou, Nuria Pelechano, Sergio Orts-Escolano, et al. 2020. The rocketbox library and the utility of freely available rigged avatars. *Frontiers in virtual reality* 1 (2020), 20.

Kylie L Goodman and Christopher B Mayhorn. 2023. It's not what you say but how you say it: Examining the influence of perceived voice assistant gender and pitch on trust and reliance. *Applied Ergonomics* 106 (2023), 103864.

Geoffrey Gorisse, Olivier Christmann, Samory Houzangbe, and Simon Richir. 2018. From robot to virtual doppelganger: Impact of avatar visual fidelity and self-esteem on perceived attractiveness. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*. 1–5.

Geoffrey Gorisse, Olivier Christmann, Samory Houzangbe, and Simon Richir. 2019. From robot to virtual doppelganger: Impact of visual fidelity of avatars controlled in third-person perspective on embodiment and behavior in immersive virtual environments. *Frontiers in Robotics and AI* 6 (2019), 8.

Rosanna E Guadagno, Jim Blascovich, Jeremy N Bailenson, and Cade McCall. 2007. Virtual humans and persuasion: The effects of agency and behavioral realism. *Media Psychology* 10, 1 (2007), 1–22.

Rosanna E Guadagno, Kimberly R Swinth, and Jim Blascovich. 2011. Social evaluations of embodied agents and avatars. *Computers in Human Behavior* 27, 6 (2011), 2380–2385.

Markus Häring, Jessica Eichberg, and Elisabeth André. 2012. Studies on grounding with gaze and pointing gestures in human-robot-interaction. In *Social Robotics: 4th International Conference, ICSR 2012, Chengdu, China, October 29-31, 2012. Proceedings 4*. Springer, 378–387.

Yuji Hatada, Shigeo Yoshida, Takuji Narumi, and Michitaka Hirose. 2019. Double shellf: What psychological effects can be caused through interaction with a doppelganger?. In *Proceedings of the 10th Augmented Human International Conference 2019*. 1–8.

Frank Hegel, Soren Krach, Tilo Kircher, Britta Wrede, and Gerhard Sagerer. 2008. Understanding social robots: A user study on anthropomorphism. In *RO-MAN 2008-the 17th IEEE international symposium on robot and human interactive communication*. IEEE, 574–579.

Darragh Higgins, Katja Zibrek, Joao Cabral, Donal Egan, and Rachel McDonnell. 2022. Sympathy for the digital: Influence of synthetic voice on affinity, social presence and empathy for photorealistic virtual humans. *Computers & Graphics* 104 (2022), 116–128.

Heinz Holling and Franzis Preckel. 2005. Self-estimates of intelligence—-methodological approaches and gender differences. *Personality and individual differences* 38, 3 (2005), 503–517.

Rosalie Hooi and Hichang Cho. 2014. Avatar-driven self-disclosure: The virtual me is the actual me. *Computers in Human Behavior* 39 (2014), 20–28.

Hanseul Jun and Jeremy Bailenson. 2020. Effects of behavioral and anthropomorphic realism on social influence with virtual humans in AR. In *2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 41–44.

Kornelius I Kammler-Sücker, Annette Löffler, Dieter Kleinböhl, and Herta Flor. 2021. Exploring virtual doppelgangers as movement models to enhance voluntary imitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29 (2021), 2173–2182.

Dominic Kao, Alejandra J Magana, and Christos Mousas. 2021a. Evaluating tutorial-based instructions for controllers in virtual reality games. *Proceedings of the ACM on human-computer interaction* 5, CHI PLAY (2021), 1–28.

Dominic Kao, Rabindra Ratan, Christos Mousas, Amogh Joshi, and Edward F Melcer. 2022. Audio matters too: How audial avatar customization enhances visual avatar customization. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–27.

Dominic Kao, Rabindra Ratan, Christos Mousas, and Alejandra J Magana. 2021b. The effects of a self-similar avatar voice in educational games. *Proceedings of the ACM on Human-Computer Interaction* 5, CHI PLAY (2021), 1–28.

Julian Paul Keenan, Aaron Nelson, Margaret O'connor, and Alvaro Pascual-Leone. 2001. Self-recognition and the right hemisphere. *Nature* 409, 6818 (2001), 305–305.

Gyoung Kim and Frank Biocca. 2018. Immersion in virtual reality can increase exercise motivation and physical performance. In *Virtual, Augmented and Mixed Reality: Applications in Health, Cultural Heritage, and Industry: 10th International Conference, VAMR 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15-20, 2018, Proceedings, Part II 10*. Springer, 94–102.

Hayeon Kim, Jinhyung Park, and In-Kwon Lee. 2023. " To Be or Not to Be Me?": Exploration of Self-Similar Effects ofAvatars on Social Virtual Reality Experiences. *IEEE Transactions on Visualization and Computer Graphics* (2023).

Kangsoo Kim, Luke Boelling, Steffen Haesler, Jeremy Bailenson, Gerd Bruder, and Greg F Welch. 2018. Does a digital assistant need a body? The influence of visual embodiment and social behavior on the perception of intelligent virtual agents in AR. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 105–114.

Kangsoo Kim, Celso M de Melo, Nahal Norouzi, Gerd Bruder, and Gregory F Welch. 2020. Reducing task load with an embodied intelligent virtual assistant for improved performance in collaborative decision making. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 529–538.

Emmanuelle P Kleinlogel, Marion Curdy, João Rodrigues, Carmen Sandi, and Marianne Schmid Mast. 2021. Doppelganger-based training: Imitating our virtual self to accelerate interpersonal skills learning. *PloS one* 16, 2 (2021), e0245960.

Claudia Krogmeier and Christos Mousas. 2020. Eye fixations and electrodermal activity during low-budget virtual reality embodiment. *Computer Animation and Virtual Worlds* 31, 4-5 (2020), e1941.

Claudia Krogmeier, Christos Mousas, and David Whittinghill. 2019. Human–virtual character interaction: Toward understanding the influence of haptic feedback. *Computer Animation and Virtual Worlds* 30, 3-4 (2019), e1883.

Justin Kruger. 1999. Lake Wobegon be gone! The" below-average effect" and the egocentric nature of comparative ability judgments. *Journal of personality and social psychology* 77, 2 (1999), 221.

Christos Kyrlitsias and Despina Michael-Grigoriou. 2022. Social interaction with agents and avatars in immersive virtual environments: A survey. *Frontiers in Virtual Reality* 2 (2022), 786665.

Luchcha Lam, Minsoo Choi, Magzhan Mukanova, Klay Hauser, Fangzheng Zhao, Richard Mayer, Christos Mousas, and Nicoletta Adamo-Villani. 2023. Effects of body type and voice pitch on perceived audio-visual correspondence and believability of virtual characters. In *ACM Symposium on Applied Perception 2023*. 1–11.

Jae-Gil Lee, Ki Joon Kim, Sangwon Lee, and Dong-Hee Shin. 2015. Can autonomous vehicles be safe and trustworthy? Effects of appearance and autonomy of unmanned driving systems. *International Journal of Human-Computer Interaction* 31, 10 (2015), 682–691.

Kwan-Min Lee and Clifford Nass. 2005. Social-psychological origins of feelings of presence: Creating social presence with machine-generated voices. *Media Psychology* 7, 1 (2005), 31–45.

Kwan Min Lee, Wei Peng, Seung-A Jin, and Chang Yan. 2006. Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human–robot interaction. *Journal of communication* 56, 4 (2006), 754–772.

Tze Wei Liew, Su-Mae Tan, Wei Ming Pang, Mohammad Tariqul Islam Khan, and Si Na Kew. 2023. I am Alexa, your virtual tutor!: The effects of Amazon Alexa's text-to-speech voice enthusiasm in a multimedia learning environment. *Education and information technologies* 28, 2 (2023), 1455–1489.

Tiina Männistö-Funk and Tanja Sihvonen. 2018. Voices from the uncanny valley: How robots and artificial intelligences talk back to us. *Digital Culture & Society* 4, 1 (2018), 45–64.

Angshuman Mazumdar and Christos Mousas. 2021. Synthesizing affective virtual reality multicharacter experiences. *Computer Animation and Virtual Worlds* 32, 3-4 (2021), e2004.

Rachel McDonnell, Martin Breidt, and Heinrich H Bülthoff. 2012. Render me real? Investigating the effect of render style on the perception of animated virtual humans. *ACM Transactions on Graphics (TOG)* 31, 4 (2012), 1–11.

R Matthew Montoya, Robert S Horton, Jack L Vevea, Martyna Citkowicz, and Elissa A Lauber. 2017. A re-examination of the mere exposure effect: The influence of repeated exposure on recognition, familiarity, and liking. *Psychological bulletin* 143, 5 (2017), 459.

Masahiro Mori. 1970. The uncanny valley: the original essay by Masahiro Mori. *IEEE Spectrum* 6 (1970).

Christos Mousas, Dimitris Anastasiou, and Ourania Spantidi. 2018. The effects of appearance and motion of virtual characters on emotional reactivity. *Computers in Human behavior* 86 (2018), 99–108.

Christos Mousas, Alexandros Koilias, Banafsheh Rekabdar, Dominic Kao, and Dimitris Anastaslou. 2021. Toward understanding the effects of virtual character appearance on avoidance movement behavior. In *2021 IEEE Virtual Reality and*

3D User Interfaces (VR). IEEE, 40–49.

Sara Moussawi and Marios Koufaris. 2019. Perceived intelligence and perceived anthropomorphism of personal intelligent agents: Scale development and validation. (2019).

Clifford Nass and Kwan Min Lee. 2000. Does computer-generated speech manifest personality? An experimental test of similarity-attraction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 329–336.

Clifford Nass, Jonathan Steuer, and Ellen R Tauber. 1994. Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 72–78.

Michael G Nelson, Alexandros Koilias, Christos-Nikolaos Anagnostopoulos, and Christos Mousas. 2022. Effects of Rendering Styles of a Virtual Character on Avoidance Movement Behavior. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 594–599.

Michael G Nelson, Alexandros Koilias, Dominic Kao, and Christos Mousas. 2023. Effects of Speed of a Collocated Virtual Walker and Proximity Toward a Static Virtual Character on Avoidance Movement Behavior. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 930–939.

Nahal Norouzi, Kangsoo Kim, Jason Hochreiter, Myungho Lee, Salam Daher, Gerd Bruder, and Greg Welch. 2018. A systematic survey of 15 years of user studies published in the intelligent virtual agents conference. In *Proceedings of the 18th international conference on intelligent virtual agents*. 17–22.

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 & Virtual Environments* 12, 5 (2003), 481–494.

Ye Pan and Anthony Steed. 2017. The impact of self-avatars on trust and collaboration in shared virtual environments. *PloS one* 12, 12 (2017), e0189078.

Dhaval Parmar, Stefan Olafsson, Dina Utami, Prasanth Murali, and Timothy Bickmore. 2022. Designing empathic virtual agents: manipulating animation, voice, rendering, and empathy to create persuasive agents. *Autonomous agents and multi-agent systems* 36, 1 (2022), 17.

Sai Krishna Pathi, Annica Kristoffersson, Andrey Kiselev, and Amy Loutfi. 2019. F-formations for social interaction in simulation using virtual agents and mobile robotic telepresence systems. *Multimodal Technologies and Interaction* 3, 4 (2019), 69.

Anna Samira Praetorius and Daniel Görlich. 2020. How avatars influence user behavior: A review on the proteus effect in virtual environments and video games. In *Proceedings of the 15th International Conference on the Foundations of Digital Games*. 1–9.

Chao Qu, Willem-Paul Brinkman, Yun Ling, Pascal Wiggers, and Ingrid Heynderickx. 2014. Conversations with a virtual human: Synthetic emotions and human responses. *Computers in Human Behavior* 34 (2014), 58–68.

Stephen Reysen. 2005. Construction of a new scale: The Reysen likability scale. *Social Behavior and Personality: an international journal* 33, 2 (2005), 201–208.

Deborah Richards, Bayan Alsharbi, and Amal Abdulrahman. 2020. Can I help you? Preferences of young adults for the age, gender and ethnicity of a Virtual Support Person based on individual differences including personality and psychological state. In *Proceedings of the Australasian Computer Science Week Multiconference*. 1–10.

Peter AM Ruijten, Antal Haans, Jaap Ham, and Cees JH Midden. 2019. Perceived human-likeness of social robots: testing the Rasch model as a method for measuring anthropomorphism. *International Journal of Social Robotics* 11 (2019), 477–494.

Maha Salem, Katharina Rohlfing, Stefan Kopp, and Frank Joublin. 2011. A friendly gesture: Investigating the effect of multimodal robot behavior in human-robot interaction. In *2011 ro-man*. IEEE, 247–252.

Franziska Schrammel, Sebastian Pannasch, Sven-Thomas Graupner, Andreas Mojzisch, and Boris M Velichkovsky. 2009. Virtual friend or threat? The effects of facial expression and gaze interaction on psychophysiological responses and emotional experience. *Psychophysiology* 46, 5 (2009), 922–931.

Ulrike Schultze. 2010. Embodiment and presence in virtual worlds: a review. *Journal of Information Technology* 25 (2010), 434–449.

Meng Ting Shih, Yi-Chieh Lee, Chih-Mao Huang, and Liwei Chan. 2023. " A feeling of déjà vu": The Effects of Avatar Appearance-Similarity on Persuasiveness in Social Virtual Reality. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–31.

Elaine Short, Justin Hart, Michelle Vu, and Brian Scassellati. 2010. No fair!! an interaction with a cheating robot. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 219–226.

Keng Siau and Weiyu Wang. 2018. Building trust in artificial intelligence, machine learning, and robotics. *Cutter business technology journal* 31, 2 (2018), 47–53.

Harrison Jesse Smith and Michael Neff. 2018. Communication behavior in embodied virtual reality. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–12.

Katta Spiel. 2021. " Why are they all obsessed with Gender?"—(Non) binary Navigations through Technological Infrastructures. In *Designing Interactive Systems Conference 2021*. 478–494.

Jan-Philipp Stein and Peter Ohler. 2017. Venturing into the uncanny valley of mind—The influence of mind attribution on the acceptance of human-like characters in a virtual reality setting. *Cognition* 160 (2017), 43–50.

Daniel Ullman, Lolanda Leite, Jonathan Phillips, Julia Kim-Cohen, and Brian Scassellati. 2014. Smart human, smarter robot: How cheating affects perceptions of social agency. In *Proceedings of the annual meeting of the cognitive science society*, Vol. 36.

Matias Volante, Sabarish V Babu, Himanshu Chaturvedi, Nathan Newsome, Elham Ebrahimi, Tania Roy, Shaundra B Daily, and Tracy Fasolino. 2016. Effects of virtual human appearance fidelity on emotion contagion in affective inter-personal simulations. *IEEE transactions on visualization and computer graphics* 22, 4 (2016), 1326–1335.

Matias Volonte, Andrew T Duchowski, and Sabarish V Babu. 2019. Effects of a virtual human appearance fidelity continuum on visual attention in virtual reality. In *Proceedings of the 19th ACM international conference on intelligent virtual agents*. 141–147.

Henriette C Van Vugt, Jeremy N Bailenson, Johan F Hoorn, and Elly A Konijn. 2008. Effects of facial similarity on user responses to embodied agents. *ACM Transactions on Computer-Human Interaction (TOCHI)* 17, 2 (2008), 1–27.

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.

Helen Wauck, Gale Lucas, Ari Shapiro, Andrew Feng, Jill Boberg, and Jonathan Gratch. 2018. Analyzing the effect of avatar self-similarity on men and women in a search and rescue game. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.

Katharina Weitz, Dominik Schiller, Ruben Schlagowski, Tobias Huber, and Elisabeth André. 2019. " Do you trust me?" Increasing user-trust by integrating virtual agents in explainable AI interaction design. In *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*. 7–9.

Evan James Williams. 1949. Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Chemistry* 2, 2 (1949), 149–168.

Makiko Yamada, Lucina Q Uddin, Hidehiko Takahashi, Yasuyuki Kimura, Keisuke Takahata, Ririko Kousa, Yoko Ikoma, Yoko Eguchi, Harumasa Takano, Hiroshi Ito, et al. 2013. Superiority illusion arises from resting-state brain networks modulated by dopamine. *Proceedings of the National Academy of Sciences* 110, 11 (2013), 4363–4367.

Nick Yee and Jeremy Bailenson. 2007. The Proteus effect: The effect of transformed self-representation on behavior. *Human communication research* 33, 3 (2007), 271–290.

Robert B Zajonc. 2001. Mere exposure: A gateway to the subliminal. *Current directions in psychological science* 10, 6 (2001), 224–228.

Katja Zibrek, Joao Cabral, and Rachel McDonnell. 2021. Does synthetic voice alter social response to a photorealistic character in virtual reality?. In *Proceedings of the 14th ACM SIGGRAPH Conference on Motion, Interaction and Games*. 1–6.

Katja Zibrek, Elena Kokkinara, and Rachel McDonnell. 2018. The effect of realistic appearance of virtual characters in immersive environments-does the character's personality play a role? *IEEE transactions on visualization and computer graphics* 24, 4 (2018), 1681–1690.

A SURVEY

We developed a survey to measure how the self-similarity of a virtual character affects human perception and user experiences. The survey comprises 42 items examining eight variables: co-presence, attentional allocation, perceived intelligence, intelligence comparison, eerie, likability, believability, perceived anthropomorphism, and one open-ended question assessing the participants' overall experience. We provide our survey along with the anchors of the scales in Table A1.

Table A1. The survey we used in our study.

#	Item	Anchors of the Scale
	Co-Presence [Biocca et al. 2001]	
Q1	I noticed the virtual character.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q2	The virtual character noticed me.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q3	The virtual character's presence was obvious to me.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q4	My presence was obvious to the virtual character.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q5	The virtual character caught my attention.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q6	I caught the virtual character's attention.	1: Very Strongly Disagree, 7: Very Strongly Agree
	Attentional Allocation [Biocca et al. 2001]	
Q7	I was easily distracted from the virtual character when other things were going on.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q8	The virtual character was easily distracted from me when other things were going on.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q9	I remained focused on the virtual character throughout our interaction.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q10	The virtual character remained focused on me throughout our interaction.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q11	The virtual character did not receive my full attention.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q12	I did not receive the virtual character's full attention.	1: Very Strongly Disagree, 7: Very Strongly Agree
	Perceived Intelligence [Moussawi and Koufaris 2019]	
Q13	The virtual character was able to operate without my intervention.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q14	The virtual character was aware of the virtual environment.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q15	The virtual character was able to set and pursue tasks by themselves in anticipation of future needs.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q16	The virtual character was able to complete tasks quickly.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q17	The virtual character was able to find and process the necessary information for completing the task.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q18	The virtual character was able to adapt/adjust their behavior based on prior events.	1: Very Strongly Disagree, 7: Very Strongly Agree
	Intelligence Comparison	
Q19	Do you think the virtual character was more intelligent than you?	1: Very Strongly Disagree, 7: Very Strongly Agree
	Eerie [Zibrek et al. 2018]	
Q20	I found the virtual character appealing.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q21	I found the virtual character eerie.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q22	I found the virtual character familiar.	1: Very Strongly Disagree, 7: Very Strongly Agree
	Likability [Reysen 2005]	
Q23	This virtual character is friendly.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q24	This virtual character is likable.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q25	This virtual character is warm.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q26	This virtual character is approachable.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q27	I would ask this virtual character for advice.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q28	I would like this virtual character as a coworker.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q29	I would like this virtual character as a roommate.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q30	I would like to be friends with this virtual character.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q31	This virtual character is physically attractive.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q32	This virtual character is similar to me.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q33	This virtual character is knowledgeable.	1: Very Strongly Disagree, 7: Very Strongly Agree
	Believability [Lam et al. 2023]	
Q34	Rate the believability of the virtual character's voice and appearance combined.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q35	Rate how well you agree with the following: I would expect the virtual character to sound like this from the way they look.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q36	Rate how well you agree with the following: I would expect the virtual character to look like this from the way they sound.	1: Very Strongly Disagree, 7: Very Strongly Agree
	Perceived Anthropomorphism [Moussawi and Koufaris 2019]	
Q37	The virtual character is able to speak like a human.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q38	The virtual character can be happy.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q39	The virtual character can be friendly.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q40	The virtual character can be respectful.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q41	The virtual character can be funny.	1: Very Strongly Disagree, 7: Very Strongly Agree
Q42	The virtual character can be caring.	1: Very Strongly Disagree, 7: Very Strongly Agree
	Overall Experience	
Q43	Please provide additional comments on your overall experience	