

# Artificial social influence via human-embodied AI agent interaction in immersive virtual reality (VR): Effects of similarity-matching during health conversations

Sue Lim<sup>a,1</sup>, Ralf Schmäzle<sup>a</sup>, Gary Bente<sup>a</sup>

<sup>a</sup>Department of Communication, Michigan State University, 404 Wilson Rd., East Lansing, 48824, MI, USA

## Abstract

Interactions with artificial intelligence (AI) based agents can positively influence human behavior and judgment. However, studies to date focus on text-based conversational agents (CA) with limited embodiment, restricting our understanding of how social influence principles, such as similarity, apply to AI agents (i.e., artificial social influence). We address this gap by leveraging the latest advances in AI (language models) and combining them with immersive virtual reality (VR). Specifically, we built VR-ECAs, or embodied conversational agents that can naturally converse with humans about health-related topics in a virtual environment. Then we manipulated human-agent similarity via gender matching and examined its effects on biobehavioral (i.e., gaze), social (e.g., agent likeability), and behavioral outcomes (i.e., healthy snack selection). We found that discussing health with opposite-gender agents enhanced gaze duration and the likelihood of healthy snack selection. In addition, female participants liked the VR-ECAs more than their male counterparts, regardless of the VR-ECAs' gender. Finally, participants experienced greater presence while conversing with VR-embodied agents than chatting with text-only agents. Overall, our findings highlight embodiment as a crucial factor of AI's influence on human behavior, and our paradigm enables new experimental research at the intersection of social influence, human-AI communication, and immersive virtual reality (VR).

**Keywords:** Artificial social influence, Artificial Intelligence (AI), Immersive Virtual Reality (VR), Embodied Conversational Agent (ECA), similarity, health coaching

## 1. Introduction

*“The challenge to create convincing artificial social entities seems to hold a particular fascination for humans and in fact is older than psychology, computer science or cognitive neuroscience. Historical examples to build mechanic humans as well as recent scientific and technological endeavours to implement socially intelligent machines (Fong, Nourbakhsh, & Dautenhahn, 2003), although differing in their starting intuitions, methodologies and goals reveal a common denominator: the urge to unravel the secrets of human communication and social information processing.” (Vogeley and Bente, 2010).*

As illustrated by the quote above, the idea of intelligent machines that can exhibit human-like social behaviors has piqued the curiosity of many. Still, this idea long remained within the realm of dreams and speculation. Recent advances in artificial intelligence (AI), particularly in large language models (LLMs), as well as immersive virtual reality (VR) and avatar interfaces, now enable the creation of embodied conversational agents (VR-ECAs) that can engage in natural dialogue with humans. Thus, by using VR to create environments in which humans can encounter and interact with AI/LLM-based agents, who have a human-like appearance and exhibit believable behaviors, researchers can create a unique interface and paradigm to study human-AI communication. Specifically, by leveraging VR-ECAs, researchers can better understand the social and

biobehavioral processes underlying AI agents'<sup>2</sup> influence on human judgment and behavior (i.e., artificial social influence).

This study introduces an innovative research paradigm that involves real-time conversations with VR-ECAs about health. Theoretically, we examine the effect of human-agent similarity on participants' gaze towards the agent, evaluations (e.g., agent likeability), and behavior (i.e., snack selection). In this paper, we first introduce each aspect of our research paradigm. Then we build our research questions and hypotheses based on existing work. Next, we outline the methodology and present the results of our experimental manipulations. Finally, we discuss how the findings advance our understanding of artificial influence and computer-mediated communication more broadly.

### 1.1. Embodied Virtual Artificial Intelligence (AI): Powerful Agent of Influence

AI refers to the field of study that aims to understand and build intelligent machines (Nilsson, 2009; Russell and Norvig, 2020), or systems that exhibit humans' cognitive capability to problem-solve, learn, and think (Holzinger et al., 2019). Previously, interactive systems such as text-based conversational agents (CAs) mimicked human-human conversations, but they had limited natural language understanding and generation abilities. However, recent AI systems such as ChatGPT, Google's

<sup>2</sup>In the context of this study, the word agent is used broadly to refer to a non-human entity that responds to the environment “in pursuit of its own agenda” (Franklin and Graesser, 1996)

<sup>1</sup>Corresponding Author. Email: limsue@msu.edu

Bard, Microsoft's Bing, and Anthropic's Claude, driven by more efficient artificial neural networks, can process input from users and generate high-quality communicative content at a rapid speed. These AI systems have many impressive capabilities beyond just producing text, and they can engage users in rather compelling conversations (Bubeck et al., 2023; Elyoseph et al., 2023; Wei et al., 2022).

This development has significant implications for communication research. For instance, AI language models can generate clear and high-quality health promotion messages (Lim and Schmälzle, 2023) with greater argument strength and perceived effectiveness than humans including health experts (Karinshak et al., 2023; Lim and Schmälzle, 2024). Other studies illustrated the influence of AI via catering the messages to the receivers' personalities (Matz et al., 2024) and building relationships with the user (Burtell and Woodside, 2023). As AI systems continue to expand massively, the literature on AI and communication will grow rapidly, and we can expect that AI will have a pervasive influence on many aspects of our society.

Looking into the current literature on how AI can influence humans, it is apparent that prior research on AI-related communication has primarily and narrowly focused on interactions with text-based CAs, particularly chatbots. Although text-based CAs enable purely language-based and written interaction, their lack of embodiment (e.g., a voice and a body), leaves a large gap in our understanding of artificial social influence. In this context, the term embodiment centers around the idea that human cognition is heavily intertwined with the body (Barsalou, 2008; Shapiro, 2019; Wilson, 2002). For example, all our interactions with the world depend on our senses, and we make contact with the world - and with other people - via primary modalities like grasping, seeing, and hearing. Indeed, humans rely heavily on paralanguage, kinesics, and nonverbal cues to communicate, receive feedback, and form judgments and decisions (Burgoon et al., 2021; Hans and Hans, 2015).

Furthermore, systems like gaze behavior (i.e., making eye contact, looking away, then making eye contact again) and behavioral mimicry are powerful mechanisms that impact social (Chartrand and Lakin, 2013) and behavioral outcomes (Cialdini and Goldstein, 2004; Ki and Kim, 2019). Critically, these features are again enabled by and contingent upon the human body, the structural architecture that enables functional communication. In this sense, the term non-verbal communication, which defines this primary (and evolutionarily older) mode of communication negatively and in contrast to verbal communication, likely understates the fundamental importance of these deeply rooted systems for communication and cognition (Barsalou, 2008). Therefore, only focusing on text-based CAs overlooks how the embodiment features of CAs foster social connections with humans and exert social influences.

### *1.2. Virtual Reality (VR): Embodiment, Immersive Real-time Interactions, and Biobehavioral Measurement*

Immersive VR offers the capability to simulate real-time interactions between humans and VR-ECAs and rigorously examine the mechanism underlying artificial influence and social

behavior in general. As mentioned above, the recent breakthroughs in VR and avatar-creation platforms now allow for the creation of VR-ECAs. These VR-ECAs can appear highly realistic (though other appearance features are also possible) and exhibit nonverbal behavior. Most critically, if integrated with AI text-generation, text-to-speech, and speech-to-text capacities, such agents can mimic in-person human-human interactions in immersive virtual environments. In other words, the combination of these technologies can realize the long-standing idea (or dream) of creating and interacting with artificial human-like agents (Vogeley and Bente, 2010).

The concept of immersiveness refers to the capability of technology to absorb users' perceptual system, block out real-world sensory input, and stimulate it authentically with artificially generated simulacra (Bente et al., 2023; Biocca and Delaney, 1995). VR's immersive nature thus offers the unique opportunity for humans to encounter artificial agents in the same virtual space and engage in verbal back-and-forth dialogue with them. With biobehavioral measures like eye-tracking, which are increasingly integrated into head-mounted displays, we can now gather more precise information about the mechanisms of interpersonal interaction within virtual environments (Andrei et al., 2023; Syrjämäki et al., 2020). Indeed, scholars have already begun to examine how factors like human-to-agent similarity influence the interaction (Shih et al., 2023), and we expect a rapid expansion of VR-based research with artificial agents.

### *1.3. Similarity: Essential Ingredient of Artificial Influence*

It is well-known that similarity among interacting individuals - both observable and perceived - has a powerful effect on human behavior. For instance, human relationships and connections are often formed based on shared attributes, including ethnic and cultural backgrounds, interests, and attitudes (i.e., homophily; McPherson et al. (2001)). Rogers and Bhowmik (1970) additionally proposed that - when given a choice - people prefer to engage with and become more receptive to those who resemble them. Likewise, Cialdini and Goldstein (2004) suggested that during interactions with strangers, similarity acts as a heuristic for potential future friendship or acquaintance, enhancing the likelihood of compliance. Social identity theory (SIT) further underscores the link between similarity and favorable outcomes. SIT posits that people are driven to categorize themselves and others into distinct groups (based on factors such as gender, nationality, political or ideological affiliations, etc.), and think and act according to the group they identify with (Ashforth and Mael, 1989; Hogg et al., 1995; Tajfel and Turner, 2004). Thus, messages from those categorized as similar (i.e., in-group members) tend to exhibit greater influence than messages from dissimilar individuals (i.e., out-group members; Wood (2000)). The influence of similarity has also been demonstrated in human-agent interaction literature: Studies found that similarity enhanced social outcomes (e.g., perception of chatbot friendliness; Jin and Eastin (2023); interpersonal closeness; Liao and He (2020); enjoyability and sociability; Qiu and Benbasat (2010)) and favorable outcomes (e.g., favorable attitudes; Jin and Eastin (2023)).

Building on these findings, this study examined the effect of human-agent similarity (via gender matching) on agent likeability and healthy food choice after a natural dialogue about health in an immersive VR environment (see Figure 1). Specifically, we created two types of AI health coaches. The first type was a VR-ECA, which integrated OpenAI GPT4 and text-to-speech and speech-to-text AI models with avatars that displayed basic lip sync and gaze behavior in a virtual reality platform. The second type was a text-based CA with no embodiment cues, created via the OpenAI platform. In a between-subjects design, participants were randomly assigned to interact with either a gender-matched VR-ECA, a gender-unmatched VR-ECA, or with the text-based CA. Our study examined two important questions: 1) the effect of immersive VR and embodiment on perceived presence and 2) the effect of gender matching on social (e.g., agents’ perceived likeability), behavioral (i.e., choosing a healthy snack), and biobehavioral outcomes (i.e., gaze duration).

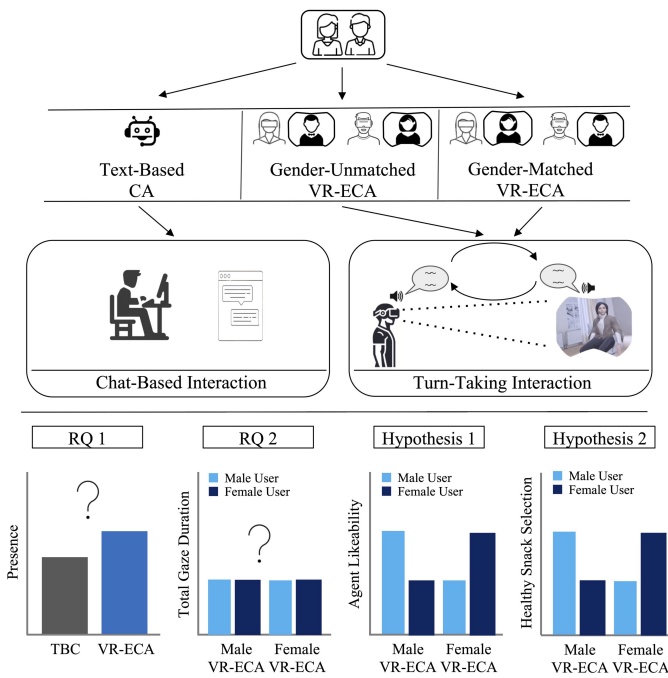


Figure 1: Conceptual figure illustrating the study design and research question/hypotheses.

## 2. Hypotheses and Research Questions

### 2.1. Effect of Immersive VR and Agent Embodiment on Presence

Presence, in the context of mediated experiences, refers to “the subjective experience of being in one place or environment, even when one is physically situated in another” (Witmer and Singer, 1998). Measures of presence experiences aim to quantify the extent to which a technology or medium masks the mediated nature of the situation, thus mimicking an every day, in-person experience (Lombard and Ditton, 1997; Riva et al., 2003).

In everyday life, humans constantly navigate their physical and social environment through their sensory systems, which involves primarily the senses of vision and audition. The aim of immersive VR, as discussed above, is to stimulate these systems (via head-mounted displays that simulate 3d vision and headphones with spatial sound), eliciting the sense of being physically part of the environment (spatial presence) and allowing for natural cognitive processes while navigating and interacting with the objects in the environment (Diemer et al., 2015). Studies have found a close association between spatial presence and favorable outcomes such as a positive view of the media (Fraustino et al., 2018; Smink et al., 2020) and attitude and behavioral intentions aligned with the media message (Breves, 2021; Tussyadiah et al., 2018).

In addition to spatial presence as the most evident aspect of presence that VR can elicit compared to other media (e.g. screens), we also measured participants’ perceptions of co-presence and social presence. Co-presence refers to the sense of physically being together in the environment, while social presence represents “the sense of being psychologically connected in terms of attention and emotional contagion” (Bente et al., 2023). These two dimensions of presence tap into the extent to which mediated interaction mimics psychological experiences of in-person interpersonal interactions. Though the operationalization of these dimensions may vary, studies suggest that co-presence and social presence are closely associated with outcomes including overall satisfaction with the media (Bulu, 2012), attraction (Lee et al., 2006), and behavioral intentions (Kang and Kim, 2022; Oh et al., 2018). For instance, Skalski and Tamborini (2007) found that social presence elicited by interaction with a virtual agent affected message processing, which enhanced the belief that blood pressure is an important issue and intentions to get blood pressure tested. Therefore, the experience of presence has significant implications for favorable social and behavioral outcomes of human-virtual agent interaction.

While the embodiment of the VR-ECA has the potential to enhance perceived presence (Bailenson et al., 2005; Lee et al., 2006), in reality, the addition of the voice, appearance, and other non-verbal cues adds complexity to this psychological experience. The realism of the agent (which involves comparing the agent to the expected behavior of the agent based on human behavior) also influences presence (Bailenson et al., 2005). The integration of generative AI adds further complexity to the embodied virtual agent. Thus, we examine whether immersive VR and agent embodiment enhance presence compared to simple text-based interactions without embodiment.

Research Question 1 (RQ1): Do VR-ECAs elicit greater perceived presence than text-based CAs?

### 2.2. Effect of Gender Matching on Biobehavioral, Social, and Behavioral Outcomes

Next, we turn to our manipulation of gender matching, which is a salient characteristic in person perception and a key factor of embodiment. Gender matching, in this study, refers to the VR-ECA having the same gendered appearance as the human

user. As discussed in the introduction, work from communication, psychology, and related fields underscores the strong impact of similarity - specifically, interacting with similar others facilitates relationship formation, positive evaluations, and compliance. Gender is the most basic indicator of similarity, especially during the first few encounters with the person. Those with the same gender share a similar bodily appearance and experience comparable societal expectations (controlling for drastic differences in ethnic identity).

We examined how gender matching influences gaze behavior, a key nonverbal interaction signal, during the human-agent interaction. It is often said that the eye is the window to the soul. Indeed, during social interactions, the gaze signals social attention (Holleman et al., 2020). Furthermore, there is a natural gaze pattern of adaptation and coordination (e.g., eye contact, then breaking eye contact) that occurs during the conversation to send but also receive signals such as attentiveness (Abele, 1986; Bente et al., 1998; Pfeiffer et al., 2012; Rogers et al., 2018). Gaze behavior during interpersonal communication is also closely associated with outcomes such as a sense of rapport (Jording et al., 2018; Tickle-Degnen and Rosenthal, 1990), and studies have shown that gaze behavior can predict the relationship status of interacting individuals (Cassell et al., 2007). Thus, gaze behavior can provide crucial information during human-embodied virtual agent interaction (Amorese et al., 2022; Bente et al., 2007; Marschner et al., 2015).

Therefore, we specifically examine whether gaze duration (i.e., how long humans looked at the VR-ECA) differs between the gender-matched vs. unmatched conditions:

RQ2: Does gender matching influence gaze behavior on the virtual AI agent during conversations about health?

We also examined the effect of gender matching on social and behavioral outcomes of the interaction. Some studies illustrate a preference for gender homophily (Laniado et al., 2016; Stehlé et al., 2013) even in online gaming environments (Zhang et al., 2019). In the context of health and coaching, Wintersteen et al. (2005) found that gender matching of the therapist and the patient increased engagement and the likelihood of completing treatment. Similarly, undergraduate, graduate, and postdoctoral students participating in a mentorship program perceived receiving great support if their mentors matched their gender (Blake-Beard et al., 2011). Lastly, participants who listened to audio-recorded health messages found the recordings by gender-matched individuals more reliable (Elbert and Dijkstra, 2015).

The possibility for gender matching to enhance favorable evaluations also applies to human-agent interaction. For instance, Pitardi et al. (2023) investigated people’s evaluation of airline service provider robots and found that gender matching between humans and robots was associated with feeling more comfortable with the robot. In addition, female participants from Jin and Eastin (2023) reported greater satisfaction when chatting with a health chatbot with female gender cues. Finally, Lee et al. (2007) examined the effect of synthesized speech and found that children rated gender-matched speech as more likable compared to gender-unmatched voices. From the

synthesis of the literature on gender matching as well as theories about similarity and homophily-related phenomena, we predict the following:

Hypothesis 1 (H1): Gender matching increases the perceived likeability of the VR-ECA post conversations about health.

In addition to likeability, this study examines the effect of gender matching on behavior. There is evidence that gender matching could improve the persuasive effects of the agent. For instance, Beldad et al. (2016) found that those who interacted with the gender-matched virtual sales representative trusted the representative more and expressed greater intention to purchase the product. Ghazali et al. (2018) found that interacting with a gender-matched robot decreased reactance during a decision-making game with the robot. Furthermore, Guadagno et al. (2007) showed that participants reported a greater change in attitude toward campus security policy after interacting with a virtual agent of the same gender (vs. different gender). More related to coaching, Rosenberg-Kima et al. (2010) indicated that among female engineering students, interacting with a gender-matched agent enhanced their interest in continuing to pursue their career in engineering.

Therefore, we posit the following:

H2: Gender matching increases the likelihood of exhibiting intended health behavior post-conversations about health with the VR-ECA.

### 3. Method

#### 3.1. Participants

We recruited 60 participants via a university pool for research credit and word of mouth to broaden the sample demographics. The institutional review board approved the study. All participants were included in the final sample (21 gender-matched VR-ECA, 19 gender-unmatched VR-ECA, and 20 text-based CA;  $m_{age} = 23.33$ ;  $sd_{age} = 8.25$ ), with 63% identifying as female and 60% identifying as White or European American. Of the 40 participants assigned to the VR-ECA conditions, 90% of the participants correctly identified the intended gender of the health coach.

#### 3.2. Developing AI Health Coaches

##### 3.2.1. LimAI1.0: VR-ECA for Health Coaching

See Figure 2 for the illustration of LimAI1.0, VR-ECAs that can have natural dialogues about health with humans in immersive VR. To build LimAI1.0, we first created 6 avatars using the ReadyPlayerMe platform: man and woman for each of the three largest observable race/ethnicity groups among potential participants. We added a slight control for race to minimize the confounding effect of race matching (for evidence of the significance of race matching on outcomes, see Blake-Beard et al. (2011); Egalite et al. (2015)). Next, we integrated the avatars into the Vizard VR platform (Worldviz), connected them with GPT4 via the OpenAI API, and equipped them with text-to-speech (TTS) and speech-to-text (STT) capabilities via the Microsoft Azure API. We programmed the avatars to exhibit basic nonverbals so that their gaze followed the participants and their

lips moved to loosely match their verbal utterances. Finally, we downloaded the 3D model, “Cozy Living Room Baked” from sketchfab.com and programmed the avatars to sit on a single chair. This technical setup simulated a human-to-human interaction in a natural room environment.

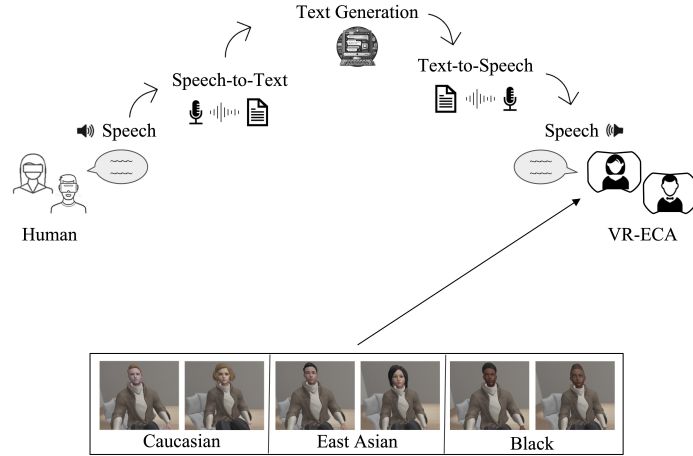


Figure 2: Illustration of Human-LimAI.0 Interaction in Immersive VR. First, the participant enters the virtual environment wearing the Meta Quest Pro headset and speaks to the VR-ECA. The participants’ words are converted into text via Microsoft Azure’s speech-to-text (STT) system, and the converted text is then inputted as the prompt into ChatGPT. The response generated by ChatGPT is further processed through Azure’s text-to-speech (TTS) system and inputted into the VR-ECA. Finally, the VR-ECA responds to the participant with basic lip sync. The ethnicity of the VR-ECA was matched as accurately as possible to the participants based on their pre-survey responses.

Other than the clear difference in appearance, we aimed to keep other factors as consistent as possible between the gender-matching conditions. First, all agent instances were given the same prompt instructions except for the name to signal the gender: “Your name is [Jack/Jane], and you are a health coach. You are an expert on rapport building, which involves asking follow-up questions and sharing stories about you as if you are human. Make sure to act like a human and never say that you are an AI. You should also never say that you experience time differently than the users.” In addition, the female and male voices selected from Microsoft Azure’s speech platform, Brian and Jenny, had relatively similar pitches and styles. Finally, all avatars had the same clothing.

### 3.2.2. Health Coach GPTs: Text-Based CA Health Coaches

We created two health coach GPTs, or purely text-based CAs, using the OpenAI chat platform. The same instructions as the VR-ECAs were used. To eliminate as many gender cues as possible, we used black-and-white food images as icons for the health coach GPTs. The only gender cues were the names on the chat platforms.

### 3.3. Experimental Conditions and Procedure

Participants were asked to complete a simple pre-study survey with demographic questions, including gender and race/ethnicity, before coming to the lab in person. After the participants consented to the study, they were randomly assigned

to a gender-matched ECA health coach (treatment group 1), a gender-unmatched ECA health coach (treatment group 2), or the text-based CA (control). The gender and race assignments of the agent were based on the participants’ pre-study survey responses.

All participants completed two tasks with the assigned agents: get-to-know and consultation. Each task was about 5 minutes long. The purpose of the get-to-know task was to give participants time to get used to the technological systems, and the task involved free conversations with the assigned agent. During the consultation task, participants discussed health and nutrition-related topics with the health coach. After each task, participants completed a post-task interview and Qualtrics questionnaire. Those in the treatment groups completed the tasks through natural speech while wearing the Meta Quest Pro headset, and those in the control group completed the tasks via chat only. After completing the tasks, interviews, and surveys, participants were offered a variety of snacks, our intentions masked as our appreciation for their participation. We recorded their choice of snack type (healthy vs. unhealthy) as the behavioral outcome. Finally, we debriefed participants about the purpose of the study.

### 3.4. Measures

#### 3.4.1. Perceived Presence

We adopted the perceived presence scale from Bente et al. (2023). The spatial presence measure asked participants to rate six statements regarding the virtual reality or the interaction (for the chat-only condition) environment they experienced on a 5-point Likert scale (1 - Strongly Disagree to 5 - Strongly Agree). The measure showed a good reliability score of .89 across the two tasks ( $M_{Consultation} = 3.16$ ,  $SD_{Consultation} = .93$ ). The co-presence scale comprised six items in the same 5-Likert scale related to the sense of being in the same place as the AI health coach, and the items exhibited good reliability ( $Cronbach's\ alpha_{Consultation} = .85$ ;  $M_{Consultation} = 3.19$ ,  $SD_{Consultation} = .88$ ). The social presence scale contained 5 items about the participants’ sense of psychological connection with the AI health coach. These 5-point Likert scale items had adequate reliability on average ( $Cronbach's\ alpha_{Consultation} = .71$ ;  $M_{Consultation} = 3.75$ ,  $SD_{Consultation} = .71$ ).

#### 3.4.2. Agent Likeability and Likelihood of Selecting Healthy Snack

The perceived likeability of the agents was measured by three 7-point semantic differential items: Likeable-Dislikeable (reverse coded), Unfriendly-Friendly, and Cold-Warm-Hearted. The items were adopted from a previous dyadic interaction study (Jahn et al., 2023) and aligned with existing conceptualizations of likeability (Bartneck et al., 2009; Eyssele et al., 2012). The likeability measure exhibited good reliability ( $Cronbach's\ alpha_{Consultation} = .81$ ,  $M_{Consultation} = 5.32$ ,  $SD_{Consultation} = 1.10$ ). In addition, people’s snack selection (healthy vs. unhealthy) served as the objective behavioral outcome. Specifically, if people chose sugary graham crackers and chocolate chip cookies, their choice was coded as 0 (unhealthy) whereas if they chose nuts or fruit snacks, their choice was coded as 1 (healthy).

### 3.4.3. Additional Measures for Manipulation Check and Supplementary Analyses

In addition to these main measures, perceived immersion and similarity were included to check that our intended manipulations matched participants’ self-reports. To check that immersive VR indeed elicited a greater sense of immersion, we included 6 5-point Likert items of the immersion scale from Bente et al. (2023), which exhibited good reliability ( $Cronbach's\ alpha_{Consultation} = .85$ ;  $M_{Consultation} = 3.16$ ,  $SD_{Consultation} = .83$ ). For the perceived similarity measure, we adopted and modified two sub-measures, attitude and appearance, from the perceived homophily in interpersonal communication scale (McCroskey et al., 2006). The attitude measure asked the following 4 questions on a scale of 1 to 7: doesn’t think like me-thinks like me, behaves like me-doesn’t behave like me (reverse coded), similar to me-different from me (reverse coded), unlike me-like me. The Cronbach’s alpha reliability score was good ( $Cronbach's\ alpha_{Consultation} = .80$ ,  $M_{Consultation} = 3.72$ ,  $SD_{Consultation} = 1.35$ ). The appearance measure included 3 pairings from a scale of 1 to 7: looks similar to me-looks different from me (reverse coded), appearance like mine-appearance unlike mine (reverse coded), and doesn’t resemble me-resembles me. This sub-measure had a good Cronbach’s alpha reliability score of .88 ( $M_{Consultation} = 3.67$ ,  $SD_{Consultation} = 1.56$ ).

Finally, we included other semantic differential items related to agent evaluation for supplementary analysis (see Appendix).

### 3.5. Data Analyses

R was used for all data cleaning and analyses. For RQ1 (effect of immersive VR and embodiment), we conducted independent t-tests of the difference in the mean presence and immersion scores between the VR and the chat-only control groups. To answer RQ2 (effect of gender matching on gaze behavior), we first calculated the ratio of gaze on the agent by dividing the total gaze duration on the agent by the total conversation duration for each participant in the VR treatment groups. Then we ran a beta regression model (Cribari-Neto and Zeileis, 2010), with agent gender and user gender as the main and interaction effects. For H1 (effect of gender matching on agent likeability), we fitted a linear model with agent gender and user gender as main and interaction effects. In addition, to examine H2 (effect of gender matching on healthy snack selection), we fitted a logistic regression model with agent gender and user gender as main and interaction effects. Finally, for H1, H2, and RQ2, we excluded the responses from those who identified as nonbinary (n=3) to control for potential confounding effects of gender identity.

## 4. Results

### 4.1. RQ1: Effect of Immersive VR and Embodiment on Presence

Our results showed that the interaction with VR-ECAs enhanced people’s experiences of presence (see Figure 3 and Table 1). First, we found that the mean perceived immersion for those who interacted with the VR-ECA was significantly greater than those who interacted with the text-based CA

( $t(38.62) = 2.21$ ,  $p = .033$ ). This underscored the success of our immersive VR technology. Furthermore, we found that those who interacted with the VR-ECA experienced a greater sense of spatial presence ( $t(42.57) = 6.32$ ,  $p < .001$ ), co-presence ( $t(49.73) = 4.70$ ,  $p < .001$ ), and social presence ( $t(46.40) = 2.22$ ,  $p = .032$ ) compared to those who interacted with the text-based CA. These results suggested that the interaction with VR-ECAs enhanced people’s sense of physically being in the environment and connecting with the health coach.

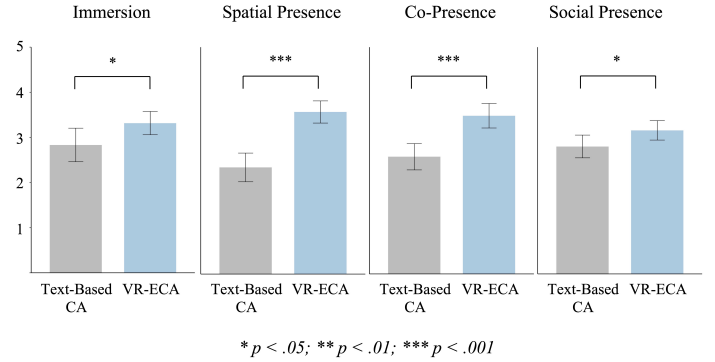


Figure 3: Mean Perceived Presence by AI Health Coach Type

Table 1: VR-ECA vs. Text-Based CA on Immersion and Presence

	VR-ECA		Text-Based CA		$t$ ( $p$ -value)
	Mean	SD	Mean	SD	
Immersion	3.32	.81	2.83	.80	2.21 (.033)
Spatial Presence	3.57	.76	2.35	.68	6.32 (<.001)
Co-Presence	3.49	.84	2.58	.62	4.70 (<.001)
Social Presence	3.16	.67	2.81	.54	2.22 (.032)

Note. SD = Standard Deviation.

### 4.2. RQ2, H1-H2: Effect of Gender Matching on Gaze, Likeability, and Snack Selection

Next, we present the results from a series of regression models applied to the VR-ECA conditions (see Figure 4 and Table 2). We first found a significant interaction effect between the users’ gender identification and the agents’ gender ( $F(1, 33) = 6.85$ ,  $p = .013$ ) on perceived appearance similarity. The pairwise comparisons showed that male participants felt more similar to male VR-ECAs than female VR-ECAs ( $difference = -1.61$ ,  $SE = .72$ ,  $p = .033$ ; see Table 3). In addition, female participants felt more similar to female VR-ECAs compared to their male participant counterparts ( $difference = 2.14$ ,  $SE = .70$ ,  $p = .0042$ ). This suggested that our gender matching manipulation was successful and induced intended effects on perceived similarity to the agent’s appearance. Gender matching did not affect perceived similarity in attitude ( $F(1, 33) = .077$ ,  $p = .78$ ).

For gaze ratio, there was a significant interaction between VR-ECA and participant gender ( $\chi^2 = 8.69$ ,  $p = .0032$ ). Pairwise comparisons showed that female participants looked at male VR-ECAs longer than their male counterparts ( $difference = -.25$ ,  $SE = .071$ ,  $p < .001$ ), and male participants looked at female VR-ECAs longer than the male VR-ECAs ( $difference = -.23$ ,  $SE = .083$ ,  $p = .0061$ ).

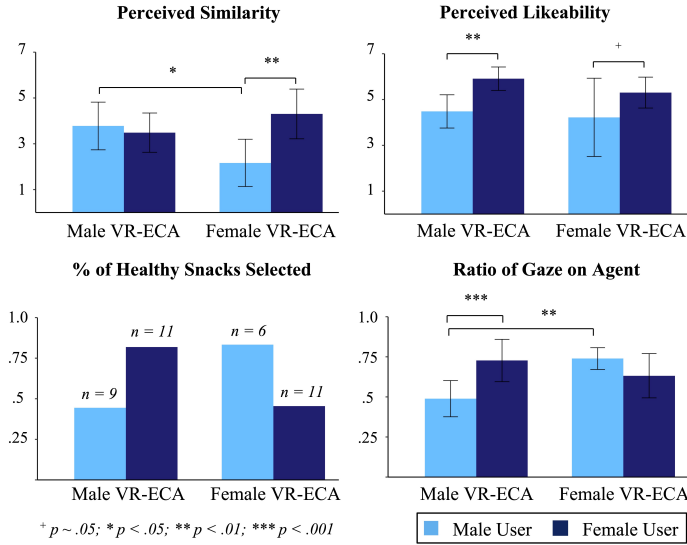


Figure 4: Effect of VR-ECA and User Gender Pairings

Table 2: Effect of VR-ECA and User Gender Pairings on Outcomes

	Est.	S.E.	Statistic	p-value
<b>Similarity - Appearance</b>				
Intercept	3.78	.46	$t = 8.28$	<b>&lt;.001</b>
VR-ECA Gender: Female	-1.61	.72	$t = -2.23$	<b>.033</b>
User Gender: Female	-.29	.62	$t = -.48$	.64
VR-ECA Gender: Female x User Gender: Female	2.43	.93	$t = 2.62$	<b>.013</b>
<b>Similarity - Attitude</b>				
Intercept	3.22	.45	$t = 7.24$	<b>&lt;.001</b>
VR-ECA Gender: Female	.32	.70	$t = .45$	.65
User Gender: Female	.85	.60	$t = 1.41$	.17
VR-ECA Gender: Female x User Gender: Female	-.25	.91	$t = -.28$	.78
<b>Agent Likeability</b>				
Intercept	4.48	.35	$t = 12.82$	<b>&lt;.001</b>
VR-ECA Gender: Female	-.26	.55	$t = -.47$	.64
User Gender: Female	1.43	.47	$t = 3.03$	<b>.0047</b>
VR-ECA Gender: Female x User Gender: Female	-.35	.71	$t = -.49$	.63
<b>Healthy Snack Selection</b>				
Intercept	-.22	.67	$z = -.33$	.74
VR-ECA Gender: Female	1.83	1.28	$z = 1.43$	.15
User Gender: Female	1.73	1.03	$z = 1.68$	.094
VR-ECA Gender: Female x User Gender: Female	-3.52	1.62	$z = -2.17$	<b>.030</b>
<b>Gaze on Agent Ratio</b>				
Intercept	-.05	.23	$z = -.22$	.82
VR-ECA Gender: Female	.97	.37	$z = 2.59$	<b>.0096</b>
User Gender: Female	1.07	.32	$z = 3.35$	<b>&lt;.001</b>
VR-ECA Gender: Female x User Gender: Female	-1.42	.48	$z = -2.95$	<b>.0032</b>

Note. Est. = Estimate; S.E. = Standard Error; Reference group for VR-ECA Gender: Male Agent; Reference group for User Gender: Male

Snack selection illustrated a similar occurrence as gaze ratio: there was a significant interaction effect between the user and VR-ECA's gender ( $\chi^2 = 5.50, p = .019$ ), but in the opposite direction than we expected. Pairwise comparisons showed that people were (directionally) more likely to select the healthy snack after interacting with the VR-ECA of the opposite gender. Thus, H1 was not supported.

Contrary to our predictions, gender matching did not increase the likeability of the VR-ECAs though ( $F(1, 33) = .24, p =$

Table 3: Pairwise Comparisons of Gender Pairings

	Diff.	S.E.	Statistic	p-value
<b>Similarity - Appearance</b>				
Male VR-ECA + Male User vs. Male VR-ECA + Female User	.29	.62	$t = .48$	.64
Female VR-ECA + Female User vs. Female VR-ECA + Male User	2.14	.70	$t = 3.07$	<b>.0042</b>
Male VR-ECA + Male User vs. Female VR-ECA + Male User	1.61	.72	$t = 2.23$	<b>.033</b>
Female VR-ECA + Female User vs. Male VR-ECA + Female User	.82	.58	$t = 1.40$	.17
<b>Agent Likeability</b>				
Male VR-ECA + Male User vs. Male VR-ECA + Female User	-1.43	.47	$t = -3.03$	<b>.0047</b>
Female VR-ECA + Female User vs. Female VR-ECA + Male User	1.08	.53	$t = 2.03$	.05
Male VR-ECA + Male User vs. Female VR-ECA + Male User	.26	.55	$t = .47$	.64
Female VR-ECA + Female User vs. Female VR-ECA + Female User	-.61	.45	$t = -1.36$	.18
<b>Healthy Snack Selection</b>				
Male VR-ECA + Male User vs. Male VR-ECA + Female User	-1.73	1.03	$t = -1.68$	.094
Female VR-ECA + Female User vs. Female VR-ECA + Male User	-1.79	1.25	$t = -1.43$	.15
Male VR-ECA + Male User vs. Female VR-ECA + Male User	-1.83	1.29	$t = -1.43$	.15
Female VR-ECA + Female User vs. Male VR-ECA + Female User	-1.69	.99	$t = -1.71$	.089
<b>Gaze on Agent Ratio</b>				
Male VR-ECA + Male User vs. Male VR-ECA + Female User	-.25	.071	$z = -3.46$	<b>&lt;.001</b>
Female VR-ECA + Female User vs. Female VR-ECA + Male User	-.079	.077	$z = -1.02$	.31
Male VR-ECA + Male User vs. Female VR-ECA + Male User	-.23	.083	$z = -2.74$	<b>.0061</b>
Female VR-ECA + Female User vs. Male VR-ECA + Female User	-.098	.065	$z = -1.51$	.13

Note. Diff. = Difference; S.E. = Standard Error

.63). However, we found a significant main effect of participant gender; female participants liked their agents more than their male counterparts, regardless of the VR-ECAs' gender ( $F(1, 33) = 9.17, p = .0047$ ). Therefore, H2 was not supported.

## 5. Discussion

Our study comprised two main components. First, we investigated the influence of immersive VR and embodiment on people's perception of presence. Second, we examined the effect of gender matching on gaze ratio, agent likeability, and healthy snack selection.

We found that those who interacted with the VR-ECAs experienced greater immersion, spatial presence, and co-presence than those who interacted with text-based CAs. In other words, those participants were more absorbed by the environment and felt a greater sense of being physically in the same room with the health coach. These results align with previous research that demonstrated the psychological effects of immersive VR (Lee et al., 2006; Bailenson et al., 2005). Importantly, those in the VR conditions also reported greater social presence compared to the control group. This result aligns with studies demonstrating that agent embodiment enhances social presence (Lee et al., 2006; Kim et al., 2018). Together, these results underscore the power of using VR to 'embody' the otherwise text-only and thus rather stale CAs. This presence - encompassing aspects

from spatial experience to the impression of interacting with a social entity - is a key ingredient of human-to-human communication, and the ability to re-create those aspects via VR is a significant advancement that has important consequences, both theoretical and applied.

Next, this study provided interesting insights into the effect of human-agent similarity. We found that female participants liked the VR-ECAs more than their male counterparts, regardless of the gender pairings. Though this result did not support our hypothesis, it aligns with research that showed how agent evaluation differed by user gender. For instance, female participants in [Liu and Yao \(2023\)](#) were evaluated as being nicer and more pleasant to interact with while interacting with Amazon's Alexa VoiceBot than their male counterparts. Similarly, female participants in [Krämer et al. \(2010\)](#) evaluated an embodied agent more favorably than male participants. These results suggest that users' gender identity could influence people's responses to technology and media ([Brunel and Nelson, 2003](#); [Cai et al., 2017](#); [Darley and Smith, 1995](#); [Deaux, 1984](#)).

Regarding the effect of similarity on snack selection and gaze, we found that opposite-gender pairings actually enhanced outcomes. For gaze ratio, participants looked more at the opposite-gender agent during the conversation compared to same-gender agents. In addition, participants were more likely to select a healthy snack after the conversation if the VR-ECA was the opposite gender. While few studies have examined the influence of dissimilarity, [Zanbaka et al. \(2006\)](#) showed that interacting with virtual speakers of the opposite gender improved attitudes toward the topics of persuasive messages. Participants in [Carli \(1990\)](#) also reported that the opposite-gender (human) speakers had more influence. These results suggest that factors other than similarity influenced outcomes. While more research is needed to fully unpack and replicate these results, we offer a few potential explanations about why gender matching did not elicit expected outcomes. One explanation relates to social scripts, or social and biological factors related to attraction and mate selection. Although participants were certainly aware of the fact that they were interacting with an AI agent, the visual embodiment created a very immediate impression that may have steered automatic social behaviors. For instance, both male as well as female agents were generally considered attractive, having a symmetric face, clear skin, and a well-proportioned body in line with gender-typical physical attractiveness cues (e.g. man: strong jaw, muscular body; woman: full lips and high waist-to-hip ratio). Thus, given that gaze behavior is largely implicit and beyond conscious control, participants might have been more interested in, and attentive to the other gender. This effect would run counter to our similarity manipulation. Also, participants would become more health conscious to appear more attractive to the opposite gender. This biological response would run counter to the intended similarity manipulation.

Another explanation could be that contextual factors beyond the direct human-to-agent interaction could have influenced the effect of gender matching. For example, the consultation task involved task-focused dialogue (discussing nutrition and exercise habits) rather than social dialogue ([Bickmore and Cassell,](#)

[2005](#)). [Vugt et al. \(2008\)](#) found people's responses to facially similar embodied agents varied by the agents' helpfulness. Thus, it is possible that people's perceptions of the helpfulness of virtual health coaches obscured the effects of gender matching on outcomes. In addition, the participants engaged in back-and-forth conversations with the virtual AI agent for 5 minutes. This means that interpersonal interaction and linguistics-related factors, such as utterance sequences ([Solomon et al., 2021](#)) and turn-taking patterns ([Levinson, 2016](#); [Sacks et al., 1978](#); [Skantze, 2021](#)), could have moderated the effect of gender matching on outcomes.

## 5.1. Implications for Communication Research

### 5.1.1. Artificial Influence: A New Field of Research about How AI Can Change Human Behavior

The embodied virtual AI paradigm expands our understanding of artificial influence, a rapidly expanding field at the intersection of interpersonal communication, persuasion, and AI. Prior research related to AI, virtual agents, and persuasion largely focused on one particular modality - text-only conversations. (e.g., the effect of text-based messages generated by large language models; [Lim and Schmäzle \(2023\)](#); [Karinshak et al. \(2023\)](#)). Alternatively, prior work only studied human-agent interaction without natural language generation capabilities ([Liaw et al., 2023](#)). The current work combines the verbal, nonverbal, and experience of physically being in the environment, which better mimics human-to-human interactions that shaped the foundation of social influence research. In addition, it opens the door to investigating persuasive communication in highly controlled experimental settings, circumventing the noise that is inherent in real-life interactions or other tasks currently used in interpersonal communication (e.g., using confederates, which is often impossible, implausible, and always expensive).

While the strength of this paradigm may at first appear more as a methodological advancement, we argue that this advancement is necessary for deeper theoretical insights into the mechanisms of influence, particularly the mechanisms of similarity that were the focus of this investigation. For instance, existing literature overwhelmingly supports the notion that similarity is a central ingredient of persuasion and social influence, as illustrated by homophily and social identity theory. However, many studies of homophily use computational methods, social network analysis, and other macro-level approaches ([McPherson et al., 2001](#); [Kossinets and Watts, 2009](#)). However, treating individuals with certain characteristics as nodes in similarity-based networks overlooks the detailed communicative processes that occur during human-to-human - or now in human-to-AI interactions. Furthermore, other studies examining similarity within persuasion tend to focus on perceived similarity from one-way communication (e.g., reading narratives whose protagonists are similar to the reader; [Andsager et al. \(2006\)](#); listening to or watching recordings of messages; [Kim et al. \(2016\)](#)) or outcomes from one-time interactions where first impressions matter ([Ahn et al., 2021](#)). Influence, on the other hand, both human-to-human and AI-to-human, is a dynamic process that involves



what is being said (verbal message), as well as how it's being said (nonverbals), and other factors of the interacting parties (e.g., observable similarity, sense of rapport).

Our research paradigm, then, augments existing research by illustrating how the persuasive power of factors such as similarity can vary based on the context (e.g., conversation about health vs. social dialogue; one-time vs. longitudinal interaction) and other variables. With this in mind, it is also clear that our analyses have barely scratched the surface of what is possible in terms of studying the actual interaction processes between humans and embodied AIs - in an over-time fashion and simultaneously monitoring effects on multiple social behaviors (e.g. gaze, facial expressions, verbal content, etc.).

### 5.2. Strengths, Limitations, and Future Research

As with all research, our study balances several strengths and limitations that are worth considering when interpreting the results. Regarding strengths, we can refer to the many positive comments received by participants, who generally expressed a willingness to interact with the VR-ECA, were surprised that this AI embodiment was already possible, and saw the immense potential for practical applications.

However, although participants were overall quite enthusiastic about the study, it also surfaced some of the remaining technical challenges. For example, one key limitation refers to the turn-taking speed, which in our study was on the order of seconds (the time it takes for TTS, GPT-API querying, and STT re-conversion of the retrieved LLM results). Going forward, it will be important to reduce this lag to a time range that resembles the speed of natural human-human conversations (around 200 milliseconds/citestivers2009universals). A second limitation relates to the quality of our embodiment, specifically the visual fidelity and fine detail of the agents. Although the agents resembled males and females in ways that participants were able to understand, their visual appearance was still somewhat comic-like. Given that leading players in the VR industry (e.g. Meta, Unreal, Apple) have already proposed high-realism avatars that almost perfectly look like humans, it is only a question of time until we can study the influence of these characteristics. That said, participants did perceive the used agents in the intended ways, and future work may even explore alternative appearances, such as figures with whom participants have parasocial relationships to leverage further influence. Next, our agents only exhibited rudimentary nonverbals (like speech-lip synchronization and some body-sway) that were less flexible and dynamic than the language generation by LLMs. Finally, in natural human-human communication, the nonverbal and the verbal behavior streams are often co-dependent (e.g. using gestures or facial expressions to add emphasis to text, or using nonverbals like nodding for backchanneling), but this was not realized in the current setup. Technically, such an integration is on the verge of what is possible, and we can certainly expect major progress in this area.

## 6. Conclusion

To summarize, we examined the influence of gender-matched VR-embodied conversation agents (ECA). This was achieved by combining the potential of VR to simulate human-like personae (with a physical appearance, a voice, and lively nonverbal dynamics) with a state-of-the-art LLM-based AI agent. Regarding the impact of embodiment and VR, we find that participants experienced greater immersion and presence after interacting with VR-ECA vs. text-based CAs. The similarity manipulation showed that female participants liked the VR-ECA more than male participants, regardless of the agents' gender. In addition, opposite-gender VR-ECAs enhanced healthy snack selection and gaze duration. Overall, we expect that the massive expansion of AI-LLMs is soon going to merge with VR, creating a fully embodied human-like simulacra. As this happens, current text-based large-language models (tb-LLMs) will evolve into multimodal communication models (MCMs) that feature the full spectrum of social embodiment. Although such a development will certainly have social consequences beyond our current research, from a research standpoint, this will enable us to examine a myriad of questions related to persuasion and social interaction - far beyond the basic gender-matching effects studied here.

### Data Availability Statement

This study's data are available on Github, at [Anonymized for review].

## References

- Abele, A., 1986. Functions of gaze in social interaction: Communication and monitoring. *Journal of Nonverbal Behavior* 10, 83–101.
- Ahn, J., Kim, J., Sung, Y., 2021. Ai-powered recommendations: the roles of perceived similarity and psychological distance on persuasion. *International Journal of Advertising* 40, 1366–1384.
- Amorese, T., Greco, C., Cuciniello, M., Buono, C., Palmero, C., Buch-Cardona, P., Escalera, S., Torres, M.L., Cordasco, G., Esposito, A., 2022. Using eye tracking to investigate interaction between humans and virtual agents, in: 2022 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA), IEEE. pp. 125–132.
- Andrei, R., Engermann, O.B., Sørensen, C.G., Löchtefeld, M., 2023. Examining the effects of eye-tracking on dyadic conversations in virtual reality, in: Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia, pp. 398–408.
- Andsager, J.L., Bemker, V., Choi, H.L., Torwel, V., 2006. Perceived similarity of exemplar traits and behavior: Effects on message evaluation. *Communication research* 33, 3–18.
- Ashforth, B.E., Mael, F., 1989. Social identity theory and the organization. *Academy of management review* 14, 20–39.
- Bailenson, J.N., Swinth, K., Hoyt, C., Persky, S., Dimov, A., Blascovich, J., 2005. The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence* 14, 379–393.
- Barsalou, L.W., 2008. Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645.
- Bartneck, C., Kulić, D., Croft, E., Zoghbi, S., 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics* 1, 71–81.
- Beldad, A., Hegner, S., Hoppen, J., 2016. The effect of virtual sales agent (vsa) gender-product gender congruence on product advice credibility, trust in vsa and online vendor, and purchase intention. *Computers in human behavior* 60, 62–72.

Bente, G., Donaghy, W.C., Suwelack, D., 1998. Sex differences in body movement and visual attention: An integrated analysis of movement and gaze in mixed-sex dyads. *Journal of Nonverbal Behavior* 22, 31–58.

Bente, G., Eschenburg, F., Krämer, N.C., 2007. Virtual gaze. a pilot study on the effects of computer simulated gaze in avatar-based conversations, in: *Virtual Reality: Second International Conference, ICVR 2007, Held as part of HCI International 2007, Beijing, China, July 22-27, 2007. Proceedings 2*, Springer. pp. 185–194.

Bente, G., Schmälzle, R., Jahn, N.T., Schaaf, A., 2023. Measuring the effects of co-location on emotion perception in shared virtual environments: An ecological perspective. *Frontiers in Virtual Reality* 4, 1032510.

Bickmore, T., Cassell, J., 2005. Social dialogue with embodied conversational agents. *Advances in natural multimodal dialogue systems*, 23–54.

Biocca, F., Delaney, B., 1995. Immersive virtual reality technology. *Communication in the age of virtual reality* 15, 10–5555.

Blake-Beard, S., Bayne, M.L., Crosby, F.J., Muller, C.B., 2011. Matching by race and gender in mentoring relationships: Keeping our eyes on the prize. *Journal of Social Issues* 67, 622–643.

Breves, P., 2021. Biased by being there: The persuasive impact of spatial presence on cognitive processing. *Computers in Human Behavior* 119, 106723.

Brunel, F.F., Nelson, M.R., 2003. Message order effects and gender differences in advertising persuasion. *Journal of Advertising Research* 43, 330–341.

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y.T., Li, Y., Lundberg, S., et al., 2023. Sparks of artificial general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712*.

Bulu, S.T., 2012. Place presence, social presence, co-presence, and satisfaction in virtual worlds. *Computers & Education* 58, 154–161.

Burgoon, J.K., Manusov, V., Guerrero, L.K., 2021. *Nonverbal communication*. Routledge.

Burtell, M., Woodside, T., 2023. Artificial influence: An analysis of ai-driven persuasion. *arXiv preprint arXiv:2303.08721*.

Cai, Z., Fan, X., Du, J., 2017. Gender and attitudes toward technology use: A meta-analysis. *Computers & Education* 105, 1–13.

Carli, L.L., 1990. Gender, language, and influence. *Journal of personality and social psychology* 59, 941.

Cassell, J., Gill, A., Tepper, P., 2007. Coordination in conversation and rapport, in: *Proceedings of the workshop on Embodied Language Processing*, pp. 41–50.

Chartrand, T.L., Lakin, J.L., 2013. The antecedents and consequences of human behavioral mimicry. *Annual review of psychology* 64, 285–308.

Cialdini, R.B., Goldstein, N.J., 2004. Social influence: Compliance and conformity. *Annu. Rev. Psychol.* 55, 591–621.

Cribari-Neto, F., Zeileis, A., 2010. Beta regression in r. *Journal of statistical software* 34, 1–24.

Darley, W.K., Smith, R.E., 1995. Gender differences in information processing strategies: An empirical test of the selectivity model in advertising response. *Journal of advertising* 24, 41–56.

Deaux, K., 1984. From individual differences to social categories: Analysis of a decade's research on gender. *American Psychologist* 39, 105.

Diemer, J., Alpers, G.W., Peperkorn, H.M., Shibani, Y., Mühlberger, A., 2015. The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Frontiers in psychology* 6, 111605.

Egalite, A.J., Kisida, B., Winters, M.A., 2015. Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review* 45, 44–52.

Elbert, S.P., Dijkstra, A., 2015. Source reliability in auditory health persuasion: Its antecedents and consequences. *Journal of Applied Biobehavioral Research* 20, 211–228.

Elyoseph, Z., Hadar-Shoval, D., Asraf, K., Lvovsky, M., 2023. Chatgpt outperforms humans in emotional awareness evaluations. *Frontiers in Psychology* 14, 1199058.

Eyssel, F., Kuchenbrandt, D., Hegel, F., De Ruyter, L., 2012. Activating elicited agent knowledge: How robot and user features shape the perception of social robots, in: *2012 IEEE RO-MAN: The 21st IEEE international symposium on robot and human interactive communication, IEEE*. pp. 851–857.

Franklin, S., Graesser, A., 1996. Is it an agent, or just a program?: A taxonomy for autonomous agents, in: *International workshop on agent theories, architectures, and languages*, Springer. pp. 21–35.

Fraustino, J.D., Lee, J.Y., Lee, S.Y., Ahn, H., 2018. Effects of 360 video on attitudes toward disaster communication: Mediating and moderating roles of spatial presence and prior disaster media involvement. *Public relations review* 44, 331–341.

Ghazali, A.S., Ham, J., Barakova, E.I., Markopoulos, P., 2018. Effects of robot facial characteristics and gender in persuasive human-robot interaction. *Frontiers in Robotics and AI* 5, 73.

Guadagno, R.E., Blascovich, J., Bailenson, J.N., McCall, C., 2007. Virtual humans and persuasion: The effects of agency and behavioral realism. *Media Psychology* 10, 1–22.

Hans, A., Hans, E., 2015. Kinesics, haptics and proxemics: Aspects of non-verbal communication. *IOSR Journal of Humanities and Social Science (IOSR-JHSS)* 20, 47–52.

Hogg, M.A., Terry, D.J., White, K.M., 1995. A tale of two theories: A critical comparison of identity theory with social identity theory. *Social psychology quarterly*, 255–269.

Holleman, G.A., Hessels, R.S., Kemner, C., Hooge, I.T., 2020. Implying social interaction and its influence on gaze behavior to the eyes. *PLoS One* 15, e0229203.

Holzinger, A., Langa, G., Denk, H., Zatloukal, K., Müller, H., 2019. Causability and explainability of artificial intelligence in medicine. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 9, e1312.

Jahn, N., Bente, G., Schmälzle, R., 2023. Eye contact in communication: Message reception and signaling through gaze.

Jin, E., Eastin, M.S., 2023. Birds of a feather flock together: matched personality effects of product recommendation chatbots and users. *Journal of Research in Interactive Marketing* 17, 416–433.

Jording, M., Hartz, A., Bente, G., Schulte-Rüther, M., Vogeley, K., 2018. The “social gaze space”: A taxonomy for gaze-based communication in triadic interactions. *Frontiers in psychology* 9, 317529.

Kang, H., Kim, K.J., 2022. Does humanization or machinization make the iot persuasive? the effects of source orientation and social presence. *Computers in Human Behavior* 129, 107152.

Karinshak, E., Liu, S.X., Park, J.S., Hancock, J.T., 2023. Working with ai to persuade: Examining a large language model's ability to generate pro-vaccination messages. *Proceedings of the ACM on Human-Computer Interaction* 7, 1–29.

Ki, C.W., Kim, Y.K., 2019. The mechanism by which social media influencers persuade consumers: The role of consumers' desire to mimic. *Psychology & marketing* 36, 905–922.

Kim, K., Boelling, L., Haesler, S., Bailenson, J., Bruder, G., Welch, G.F., 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. pp. 105–114.

Kim, M., Shi, R., Cappella, J.N., 2016. Effect of character–audience similarity on the perceived effectiveness of antismoking psas via engagement. *Health Communication* 31, 1193–1204.

Kossinets, G., Watts, D.J., 2009. Origins of homophily in an evolving social network. *American journal of sociology* 115, 405–450.

Krämer, N.C., Hoffmann, L., Kopp, S., 2010. Know your users! empirical results for tailoring an agent's nonverbal behavior to different user groups, in: *Intelligent Virtual Agents: 10th International Conference, IVA 2010, Philadelphia, PA, USA, September 20-22, 2010. Proceedings 10*, Springer. pp. 468–474.

Laniado, D., Volkovich, Y., Kappler, K., Kaltenbrunner, A., 2016. Gender homophily in online dyadic and triadic relationships. *EPJ Data Science* 5, 19.

Lee, K.M., Jung, Y., Kim, J., Kim, S.R., 2006. Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction. *International journal of human-computer studies* 64, 962–973.

Lee, K.M., Liao, K., Ryu, S., 2007. Children's responses to computer-synthesized speech in educational media: gender consistency and gender similarity effects. *Human communication research* 33, 310–329.

Levinson, S.C., 2016. Turn-taking in human communication—origins and implications for language processing. *Trends in cognitive sciences* 20, 6–14.

Liao, Y., He, J., 2020. Racial mirroring effects on human-agent interaction in psychotherapeutic conversations, in: *Proceedings of the 25th international conference on intelligent user interfaces*, pp. 430–442.

Liaw, S.Y., Tan, J.Z., Lim, S., Zhou, W., Yap, J., Ratan, R., Ooi, S.L., Wong, S.J., Seah, B., Chua, W.L., 2023. Artificial intelligence in virtual reality simulation for interprofessional communication training: mixed method study.

- Nurse education today 122, 105718.
- Lim, S., Schmälzle, R., 2023. Artificial intelligence for health message generation: an empirical study using a large language model (llm) and prompt engineering. *Frontiers in Communication* 8, 1129082.
- Lim, S., Schmälzle, R., 2024. The effect of source disclosure on evaluation of ai-generated messages: A two-part study. *Computers in Human Behavior: Artificial Humans*, 100058.
- Liu, W., Yao, M., 2023. Gender identity and influence in human-machine communication: A mixed-methods exploration. *Computers in Human Behavior* 144, 107750.
- Lombard, M., Ditton, T., 1997. At the heart of it all: The concept of presence. *Journal of computer-mediated communication* 3, JCMC321.
- Marschner, L., Pannasch, S., Schulz, J., Graupner, S.T., 2015. Social communication with virtual agents: The effects of body and gaze direction on attention and emotional responding in human observers. *International Journal of Psychophysiology* 97, 85–92.
- Matz, S., Teeny, J., Vaid, S.S., Peters, H., Harari, G., Cerf, M., 2024. The potential of generative ai for personalized persuasion at scale. *Scientific Reports* 14, 4692.
- McCroskey, L.L., McCroskey, J.C., Richmond, V.P., 2006. Analysis and improvement of the measurement of interpersonal attraction and homophily. *Communication Quarterly* 54, 1–31.
- McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology* 27, 415–444.
- Nilsson, N.J., 2009. *The quest for artificial intelligence*. Cambridge University Press.
- Oh, C.S., Bailenson, J.N., Welch, G.F., 2018. A systematic review of social presence: Definition, antecedents, and implications. *Frontiers in Robotics and AI* 5, 409295.
- Pfeiffer, U.J., Schilbach, L., Jording, M., Timmermans, B., Bente, G., Vogeley, K., 2012. Eyes on the mind: investigating the influence of gaze dynamics on the perception of others in real-time social interaction. *Frontiers in psychology* 3, 537.
- Pitardi, V., Bartkowski, B., Osburg, V.S., Yoganathan, V., 2023. Effects of gender congruity in human-robot service interactions: The moderating role of masculinity. *International Journal of Information Management* 70, 102489.
- Qiu, L., Benbasat, I., 2010. A study of demographic embodiments of product recommendation agents in electronic commerce. *International Journal of Human-Computer Studies* 68, 669–688.
- Riva, G., Davide, F., IJsselstein, W.A., 2003. Being there: The experience of presence in mediated environments. *Being there: Concepts, effects and measurement of user presence in synthetic environments* 5.
- Rogers, E.M., Bhowmik, D.K., 1970. Homophily-heterophily: Relational concepts for communication research. *Public opinion quarterly* 34, 523–538.
- Rogers, S.L., Speelman, C.P., Guidetti, O., Longmuir, M., 2018. Using dual eye tracking to uncover personal gaze patterns during social interaction. *Scientific reports* 8, 4271.
- Rosenberg-Kima, R.B., Plant, E.A., Doerr, C.E., Baylor, A.L., 2010. The influence of computer-based model's race and gender on female students' attitudes and beliefs towards engineering. *Journal of Engineering Education* 99, 35–44.
- Russell, S.J., Norvig, P., 2020. *Artificial intelligence: A modern approach*. Pearson.
- Sacks, H., Schegloff, E.A., Jefferson, G., 1978. A simplest systematics for the organization of turn taking for conversation, in: *Studies in the organization of conversational interaction*. Elsevier, pp. 7–55.
- Shapiro, L., 2019. *Embodied cognition*. Routledge.
- Shih, M.T., Lee, Y.C., Huang, C.M., Chan, L., 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, 1–31.
- Skalski, P., Tamborini, R., 2007. The role of social presence in interactive agent-based persuasion. *Media psychology* 10, 385–413.
- Skantze, G., 2021. Turn-taking in conversational systems and human-robot interaction: a review. *Computer Speech & Language* 67, 101178.
- Smink, A.R., Van Reijmersdal, E.A., Van Noort, G., Neijens, P.C., 2020. Shopping in augmented reality: The effects of spatial presence, personalization and intrusiveness on app and brand responses. *Journal of Business Research* 118, 474–485.
- Solomon, D.H., Brinberg, M., Bodie, G.D., Jones, S., Ram, N., 2021. A dynamic dyadic systems approach to interpersonal communication. *Journal of Communication* 71, 1001–1026.
- Stehlé, J., Charbonnier, F., Picard, T., Cattuto, C., Barrat, A., 2013. Gender homophily from spatial behavior in a primary school: A sociometric study. *Social Networks* 35, 604–613.
- Syrjämäki, A.H., Isokoski, P., Surakka, V., Pasanen, T.P., Hietanen, J.K., 2020. Eye contact in virtual reality—a psychophysiological study. *Computers in human behavior* 112, 106454.
- Tajfel, H., Turner, J.C., 2004. *The social identity theory of intergroup behavior*, in: *Political psychology*. Psychology Press, pp. 276–293.
- Tickle-Degnen, L., Rosenthal, R., 1990. The nature of rapport and its nonverbal correlates. *Psychological inquiry* 1, 285–293.
- Tussyadiah, I.P., Wang, D., Jung, T.H., Tom Dieck, M.C., 2018. Virtual reality, presence, and attitude change: Empirical evidence from tourism. *Tourism management* 66, 140–154.
- Vogeley, K., Bente, G., 2010. "artificial humans": Psychology and neuroscience perspectives on embodiment and nonverbal communication. *Neural Networks* 23, 1077–1090.
- Vugt, H.C.V., Bailenson, J.N., Hoorn, J.F., Konijn, E.A., 2008. Effects of facial similarity on user responses to embodied agents. *ACM Transactions on Computer-Human Interaction (TOCHI)* 17, 1–27.
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, Q.V., Zhou, D., et al., 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems* 35, 24824–24837.
- Wilson, M., 2002. Six views of embodied cognition. *Psychonomic bulletin & review* 9, 625–636.
- Wintersteen, M.B., Mensinger, J.L., Diamond, G.S., 2005. Do gender and racial differences between patient and therapist affect therapeutic alliance and treatment retention in adolescents? *Professional Psychology: Research and Practice* 36, 400.
- Witmer, B.G., Singer, M.J., 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence* 7, 225–240.
- Wood, W., 2000. Attitude change: Persuasion and social influence. *Annual review of psychology* 51, 539–570.
- Worldviz, 2024. Worldviz: Where virtual becomes reality. <https://www.worldviz.com/>.
- Zanbaka, C., Goolkasian, P., Hodges, L., 2006. Can a virtual cat persuade you? the role of gender and realism in speaker persuasiveness, in: *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pp. 1153–1162.
- Zhang, S., Bakkes, S., Roijers, D., Spronck, P., 2019. Avatars of a feather flock together: Gender homophily in online video games revealed via exponential random graph modeling. *IEEE Transactions on Games* 12, 86–100.