Artificial Social Influence: Rapport-Building, LLM-Based Embodied Conversational Agents for Health Coaching

Sue Lim[1,*], Ralf Schmälzle[1] and Gary Bente[1]

[1]CARISMA Lab, Michigan State University, 404 Wilson Rd. East Lansing, Michigan, USA

Abstract

Embodied conversational agents (ECAs) have been designed and implemented to provide support to humans, especially in the area of health. With the recent advancements in large language models (LLMs), ECAs can now be equipped with natural language capabilities, engaging in turn-taking dialogue with humans while exhibiting verbal rapport-building behavior. Our innovative study designed LLM-based ECAs that provide health coaching to people through immersive virtual reality (VR). Specifically, male and female avatars were integrated with ChatGPT, text-to-speech, and speech-to-text APIs on a VR platform. For our experiment, we manipulated human-ECA similarity via gender-matching. Participants were randomly assigned to either a gender-matched or unmatched embodied health coach and completed two interaction tasks (get-to-know-you and health consultation) in immersive VR. Our quantitative evaluations showed that those in the gender-unmatched conditions rated certain interaction metrics more favorably compared to those in the gender-matched condition. The qualitative evaluation showed that while the lack of nonverbals and other technology-related limitations could be improved, the LLM-based ECAs showed the potential to support people's health-related decision-making.

Keywords

Artificial Intelligence (AI), Embodied Conversational Agents (ECAs), Virtual Reality (VR), Artificial Social Influence, Health Coaching

1. Introduction and Study Overview

Embodied conversational agents (ECAs), or artificial entities that can interact face-to-face with humans through verbal and non-verbal cues [1, 2, 3], have been part of human society for decades. Research shows that ECAs can build relationships with humans through realistic human-like behavior such as gaze, head movement, and smiling during conversations [4, 5]. As a result, ECAs have been designed and implemented as pedagogical, training, and health support tools [6, 7, 8]. Despite ECAs' potential to naturally interact with humans and influence human behavior, ECAs thus far exhibited limited natural language capabilities: Most used rule-based or decision-tree dialogue systems based on the user's selection from pre-determined responses or wizard-of-oz paradigms featuring a human behind the agent. Since many human communication processes, such as social influence and persuasion, occur through language exchange, this limitation left a significant gap in research on human-ECA communication.

The recent advancements in artificial intelligence (AI), especially state-of-the-art large language models (LLMs) such as OpenAI's GPT4, have significantly improved this limitation. Now, AI-driven CAs can engage people in natural dialogue while expressing verbal rapport-building behavior [9], including empathy [10] and active listening [11]. ECAs' verbal rapport-building capabilities have significant implications on *artificial social influence*, or AI-driven agents' influence on human thought, emotions, and behavior. First, extant literature underscores the importance of rapport[1] for successful human social interactions [16]. In fact, rapport is strongly linked with relational outcomes, including student performance (teacher-student interaction [12, 17, 18]), patient health and adherence to treatment (provider-patient interaction [19, 20, 21]), and social influence and persuasion more broadly (communicator-receiver interaction [22]). Scholars in the intelligent virtual agents (IVA) community have already begun examining rapport, mostly in the context of agents' nonverbal behavior [13, 23, 24]

Therefore, we built upon existing work in IVA and human-computer interaction more broadly by designing and evaluating LLM-based ECAs who facilitate healthy decision-making by engaging people in natural dialogue while exhibiting basic nonverbal cues in immersive virtual reality (VR). We specifically focused on nutrition and physical activity for this proof-of-concept study because of its relevance to any demographic. Grounded in

CONNECT Workshop: 24th ACM International Conference on Intelligent Virtual Agents, Sep. 16-19, 2024

*Corresponding author.
limsue@msu.edu (S. Lim); schmaelz@msu.edu (R. Schmälzle); gabente@msu.edu (G. Bente)
0000-0001-7559-1317 (S. Lim); 0000-0002-0179-1364 (R. Schmälzle); 0009-0005-0819-8132 (G. Bente)

CEUR Workshop Proceedings (CEUR-WS.org)

[1]In the context of this study, we define rapport as a harmonious relational dynamic that fosters open dialogue, a cooperative atmosphere, and a sense of mutual social connectedness, respect, and trust among its members [12, 13, 14, 15].

literature about the link between similarity and social influence outcomes [22, 25, 26], we manipulated human-ECA similarity via gender-matching (a basic factor of similarity) and had participants complete two interaction tasks, get-to-know-you and health consultation, with the assigned agent.

The preliminary results found that people who interacted with the gender-unmatched agent showed a slight trend toward selecting a healthier snack than those who interacted with the gender-matched agent [27][2]. This follow-up paper takes a deep dive into the development of the LLM-driven health coaches as well as the evaluation of the interactions from the participants' perspective. Specifically, we examine the following general research questions (RQs):

RQ1: Does human-ECA similarity influence people's experiences during the interaction? **RQ2**: Does human-ECA similarity influence people's evaluation of the interaction? **RQ3**: To what extent are people satisfied with the LLM-based embodied health coaches?

2. LLM-Based ECA Design and Development

The design and development process included three stages: 1) avatar creation in a VR platform [3], 2) integration of the LLM and prompting it to become health coaches, and 3) adding the speech-to-text and text-to-speech systems to the avatars. First, we created six characters, one male and one female, for three main racial groups commonly seen on our university's campus on the ReadyPlayerMe platform [4] (see Figure 1). We did this to control for the potential effects of racial matching on our outcomes. Then, we updated the avatars on the VR platform and programmed them to exhibit basic nonverbal cues. Specifically, we programmed the agents to keep their gaze on the participant at all times and display simple lip sync when speaking.

For the LLM, we used OpenAI's API to leverage GPT4 LLM[5]. We named the male health coach Jack and the female health coach Jane, common names for each gender. Next, we provided the following instructions to GPT4: "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." The instructions were finalized through the prompt engineering process. During this process, the authors went through iterations of putting in the instructions, conversing with the health coach, and adjusting the instructions to make the conversation sound more natural and believable. Through prompt engineering, we found that "you are a health coach" was enough for the health coach to respond with appropriate health information and discuss making healthy decisions. However, a longer instruction about rapport and response style had to be added to elicit relational language in the health coach's response.

Afterward, we integrated the text-to-speech (TTS) and speech-to-text (STT) services via the Microsoft Azure API[6]. We chose Brian as Jack's voice and Jenny as Jane's voice because the voices had relatively similar pitches and styles. For the virtual office setting, we downloaded the 3D model called "Cozy Living Room Baked" from Sketchfab[7] and uploaded it to Vizard. Finally, we had the agents sit on a single chair in the room, their bodies angled toward the couch. The participants were set up to "sit" on the couch when they entered the virtual environment.

3. Methodology

In this section, we provide descriptions about our participants, outline the experimental procedures, and explain the quantitative and qualitative measures used to answer our research questions.

The participants for the study consisted of 42 individuals recruited [8] 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 ECAs and 21 gender-unmatched ECAs; M_{age} = 23.2; SD_{age} = 9.3).

3.1. Experimental Procedures

Before coming into the lab, participants completed a pre-survey and provided information about their gender and

[2]This first paper from our study explicitly focused on the health consultation task and outcomes related to agent evaluation, biobehavioral outcome, and health-related decision-making. It also featured a discussion comparing the ECAs to the text-based conversational agent system. Please see this paper for a thorough literature review on the importance of similarity within human-computer interaction as well as human-to-human interaction.

[3]Worldviz (2023). We build VR labs. https://www.worldviz.com.

[4]Ready Player Me (2024). Integrate customizable avatars into your game or app in minutes. https://readyplayer.me/

[5]OpenAI (2022). Introducing ChatGPT. https://openai.com/index/chatgpt/

[6]Microsoft (2024). Azure AI Speech. https://azure.microsoft.com/en-us/products/ai-services/ai-speech

[7]Cozy living room baked (2022). SketchFab. https://sketchfab.com/3d-models/cozy-living-room-baked-581238dc5fda4dc990571cdc02827783

[8]Two of the 42 participants only completed the qualitative evaluations. Thus, they are not included in the quantitative metrics

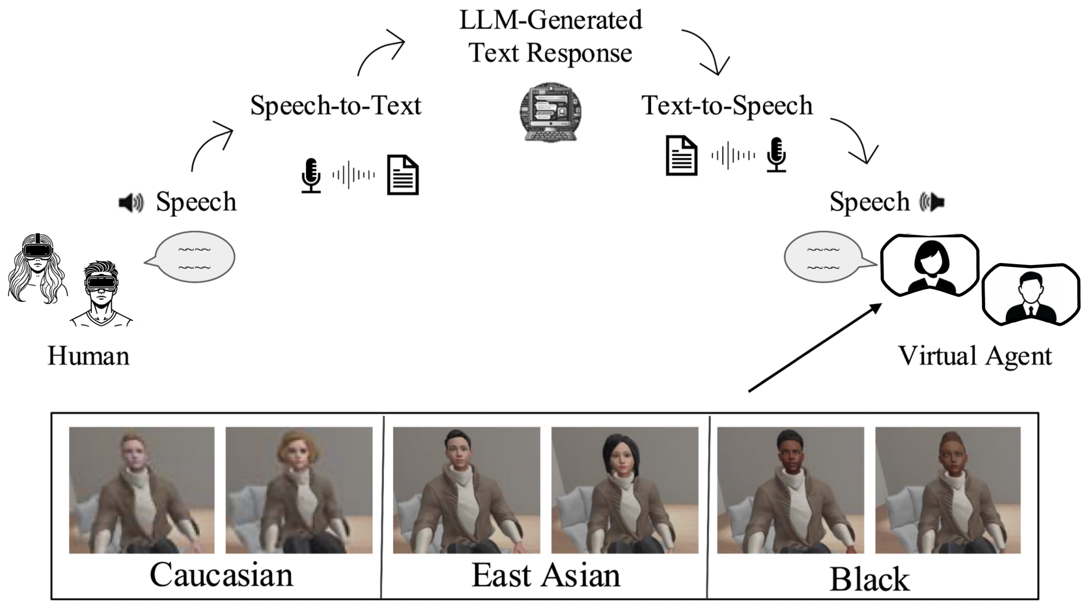

Figure 1: Illustration of the AI Components Underlying Human-ECA Interaction: First, the participant puts on the VR headset and speaks. The words are converted into text via a speech model and inputted into ChatGPT as the prompt. ChatGPT generates response, and the health coach speaks the words with basic lip sync and eye contact that always follows the participants. ECA's racial group was assigned to match the participant's reported racial group.

race identification. Once they consented[9] to the study, we randomly assigned them to either the gender-matched or unmatched conditions based on the pre-survey responses. Then, they put on the Meta Quest VR headset and completed two 5-minute interaction tasks: get-to-know-you and health consultation. Get-to-know-you involved having casual conversations about any topic with the health coach. The health consultation task required a focused conversation about nutrition, physical activity, or general well-being (see Appendix for task instruction details). After each task, participants completed a brief semi-structured interview and Qualtrics questionnaire about their experience. At the end the study, we offered a variety of snacks to the participants to thank them for their participation. In reality, their choice of snack type (healthy vs. unhealthy) was noted as the behavioral outcome of the interaction. Lastly, we fully debriefed the participants about the purpose of the study.

3.2. Quantitative Measures

To examine the first two research questions about the effect of similarity, we asked participants to complete a post-task survey about the interaction, self-perception, and co-presence after each interaction task. Specifically, **the interaction and self-perception-related items** asked people to rate on a 7-point Likert scale (1 - Strongly disagree, 7 - Strongly agree) about seven rapport-related factors of the interaction (i.e., was friendly, was warm, was satisfying, was harmonious, had rapport, had focus, ran smoothly) and four satisfaction related factors (i.e., I was involved with the task, I enjoyed the conversation, I felt comfortable, I was satisfied with the outcome).

Furthermore, for **co-presence**, we adopted the co-presence scale from Bente et al. [28]. The co-presence scales included six 5-point Likert scale items (1 - Strongly disagree, 5 - Strongly agree) related to the sense of being in the same place as the ECA: It felt as though the avatar was with me in the room; It felt like I was in the same space as the avatar; It felt as if the avatar and I were together in the same space; I felt that I could approach the avatar; I felt like an external observer to the scene in the virtual environment; It felt like I could physically meet the avatar in the virtual environment. We included the co-presence scale because it has been included in previous studies to examine experienced rapport during interactions with ECAs (e.g., [29]).

[9]Participants were clearly told that they were talking with an AI-driven health coach.

Table 1
Set Interview Questions for the Get-to-Know-You Task

How did you experience the interaction?
Did anything stand out to you?
How familiar are you with ChatGPT and AI in general?

Table 2
Set Interview Questions for the Consultation Task

How did you experience the interaction?
Did anything stand out to you?
What did you discuss with the health coach?
How helpful/effective did you find the conversation?
Would you talk to the health coach again (scale of 1-10)?
What else would you like to discuss with them?

3.3. Qualitative Measures

To examine RQ3, we conducted a brief semi-structured interview after each task. During the interviews, we asked them set questions (see Tables 1 and 2) and then asked follow-up questions when needed to gain deeper insights into the participants' experience.

The interviews were recorded with the participants' permission. For the data analysis, the first author transcribed the recordings and then conducted a thematic analysis [30] via line-by-line coding in Nvivo 14. Though the author did not have specific categories in mind while coding, the themes were related to the participants' positive and negative feedback about the ECAs' conversational and relational capabilities, their thoughts about the information provided by the ECA, and factors driving the participants' willingness to engage with the LLM-based ECAs.

4. Results

The quantitative (examining RQ1 and RQ2) and qualitative (examining RQ3) evaluations provided interesting insights into the effect of similarity on people's experiences and overall satisfaction with the LLM-driven health coaches.

4.1. Quantitative Evaluation

We conducted a series of independent t-tests (α =.05) in R, comparing the differences in the outcome measures by gender-matching. The results showed that across the metrics, there were not many significant differences between the gender-matched and unmatched conditions (see Table 3). For some of the metrics, we found that those in the gender-unmatched condition actually provided higher ratings than those in the gender-matched condition. For instance, the evaluation of the get-to-know-you task showed that those in the gender-unmatched condition found the interaction more harmonious (*t[35.48]* = 2.74, p =.0094), experienced greater co-presence (*t[30.57]* = 2.44, p =.021), and showed a trend toward higher satisfaction with the outcome (*t[32.09]* = 2.03, p =.051). Evaluation of the consultation task showed similar results, with the gender-unmatched condition finding the interaction more harmonious (*t[36.35]* = 2.28, p =.028; see Table 4).

4.2. Qualitative Evaluation

Upon first interacting with the health coach, 7 participants were surprised by the depth of the health coaches' responses, and 1 participant shared that they kept asking personal questions, forgetting that the health coach was an AI. Three other participants appreciated that the health coach demonstrated knowledge about topics such as Leonardo Davinci's paintings or a particular books the participants were reading. One participant shared that seeing and hearing the health coaches' responses made the information come more to life.

In addition, across the two tasks, 8 participants discussed how the health coach captured and remembered everything they said during the conversations and responded appropriately. For example, a participant with actual experience meeting with a nutritionist mentioned that they tended to forget small details, whereas the health coach responded to each point the participant said. This led to enhanced perceptions of attentiveness. Furthermore, the health coaches' verbal rapport-building behaviors such as asking questions and sharing relevant stories made the interaction feel like a real conversation for 7 participants. In fact, 15 participants thought the interaction was realistic and natural, similar to their interactions with other people. One participant even shared that during the health consultation task, they "started to answer questions in like more informal way because I kind of already know her."

On the other hand, others experienced more difficulties in the conversations. Sixteen participants discussed the difficulties of the lag in the health coaches' responses or the audio cutting in and out. Three participants shared that sometimes the responses did not sound natural enough, and it felt as though someone was reading from the internet or generated message rather than engaging in a conversation. The discussion about nonverbal cues also emerged frequently - 10 people mentioned how the mismatch between the lip sync and the words made it difficult for them to know when the health coach was done talking because of misaligned timing. Furthermore, health coaches' gaze that followed the participants' movements and didn't break eye contact elicited discom-

Table 3
Gender Matched vs. Unmatched: Get-to-Know-You Task

Metric	Unmatched Mean(SD)	Matched Mean(SD)	Effect Size (d)	*t(p-value)*
Interaction Focused	6.19 (.68)	5.68 (1.11)	.55	1.72 (.096)
Interaction Ran Smoothly	5.29 (1.35)	4.58 (1.26)	.54	1.71 (.095)
Interaction Warm	5.05 (1.46)	4.89 (1.37)	.11	.34 (.735)
Interaction Friendly	6.05 (.59)	5.84 (.60)	.34	1.09 (.283)
Interaction Harmonious	5.43 (1.16)	4.32 (1.38)	.87	2.74 (**.009**)
Interaction Satisfying	5.19 (1.12)	4.63 (1.30)	.46	1.45 (.156)
Had Good Rapport	5.43 (1.16)	4.79 (1.44)	.49	1.54 (.134)
Satisfied with Outcome	5.86 (1.01)	5.05 (1.43)	.65	2.03 (*.051*)
Involved with Task	6.29 (.72)	6.37 (.60)	.13	-.40 (.693)
Felt Comfortable	5.52 (1.36)	5.58 (1.02)	.046	-.15 (.885)
Enjoyed Conversation	5.67 (1.02)	5.26 (1.45)	.32	1.01 (.320)
Perceived Co-Presence	3.80 (.55)	3.25 (.84)	.78	2.44 (**.021**)

Table 4
Gender Matched vs. Unmatched: Health Consultation Task

Metric	Unmatched Mean(SD)	Matched Mean(SD)	Effect Size (d)	*t(p-value)*
Interaction Focused	5.86 (1.11)	5.68 (1.25)	.55	.46 (.648)
Interaction Ran Smoothly	5.48 (1.25)	5.11 (1.41)	.54	.88 (.386)
Interaction Warm	5.05 (1.50)	4.58 (1.22)	.11	1.09 (.283)
Interaction Friendly	5.76 (1.26)	5.21 (1.08)	.34	1.49 (.145)
Interaction Harmonious	5.48 (1.17)	4.58 (1.30)	.72	2.28 (**.028**)
Interaction Satisfying	5.52 (1.17)	5.05 (1.27)	.39	1.22 (.231)
Had Good Rapport	5.33 (1.24)	4.95 (1.27)	.31	.97 (.337)
Satisfied with Outcome	6.00 (1.14)	5.42 (.90)	.56	1.79 (.082)
Involved with Task	6.00 (1.30)	6.16 (.60)	.16	-.50 (.621)
Felt Comfortable	5.62 (1.07)	5.26 (.99)	.34	1.09 (.282)
Enjoyed Conversation	5.67 (1.16)	5.21 (1.44)	.35	1.10 (.279)
Perceived Co-Presence	3.72 (.84)	3.23 (.79)	.61	1.92 (.063)

fort for 6 participants. Health coaches' lack of dynamic nonverbal feedback made the conversation awkward and made the LLM-driven ECA appear robotic or cold (8 participants).

Regarding the health information provided by the health coach, 32 acknowledged that the information was helpful and the health coach appeared knowledgeable. However, the level of perceived effectiveness of the information varied. For example, participants who asked specific questions relevant to their current health (e.g., meal prep ideas, incorporating more protein to keto diet, effect of yoga) found the information effective. However, 8 participants expressed they already knew the information from the health coach and 7 participants found the information generic, something they could get over the internet. One participant specifically mentioned that the health coach lacked the ability to understand the nuances of his questions about exercise.

The next theme related to factors that influence people's willingness to engage with the LLM-based embodied health coach again. On a scale of 1 (never want to talk to the health coach again) to 10 (would want to talk to it the next day), participants, on average, leaned toward interacting with it again (M=6.2, SD = 2.4[10]).

Reasons for wanting to talk to the health coach again included health coach's availability, the novelty of the technology, knowledge about various topics, and pleasantness. A participant from Brazil discussed that it is difficult for her to connect with her family at times because of the time difference; thus, it would benefit her to have a health coach who can listen to her and support her in times of need (gave rating 10). Another participant mentioned that it is hard to discuss personal problems, even with close family members, because it could alter their judgement of the participant. Thus, the participant would want to share with the health coach and ask for advice because the health coach will not judge them.

Reasons for not wanting to talk to the health coach again included the lag in the text-to-speech system and

[10]There were no differences between experimental groups. Six participants could not provide an exact number when answering the question

the lack of nonverbal cues. Four participants stated that it would be more convenient to get information on google. Two participants acknowledged the benefit of talking with the health coach, but preferred to discuss urgent and more specific topics with another human.

5. Discussion

Overall, this proof-of-concept study provided interesting insights into participants' interaction with LLM-driven ECA in immersive VR. Through quantitative measures, we found that people's experiences related to rapport and evaluation of the interaction were generally consistent across similarity matching conditions **(RQ1-2)**. In fact, those who talked with the gender-unmatched health coach found the interaction as more harmonious (for both tasks) and perceived higher level of co-presence (for the get-to-know-you task). We suggest some potential explanations for this result in the limitations section below.

Through the qualitative interview, we found that participants leaned, on average, toward engaging with the LLM-driven health coach again **(RQ3)**. Though there was general consensus that the information provided by the health coach was accurate and helpful to some extent, the findings showed variance in expressed satisfaction with the health coach. With some participants, the health coaches' verbal rapport-building behaviors (e.g., connecting through self-disclosure, asking questions) and ability to remember the details of the conversations enhanced engagement and perceived care. However, for others, the lack of dynamic nonverbal feedback and the mismatch between lip movements and speech, among other reasons, prevented deeper involvement in the conversation. This emphasis on nonverbal behavior's role in rapport-building aligns with extant literature (e.g., [23, 14].

5.1. Limitations and Future Directions

Like all research, the current work has several limitations. One limitation concerns the strength of the experimental manipulation (gender-matching) and the resulting sample sizes needed to demonstrate possible effects. Specifically, it could be argued that the gender-matching manipulation was neither highly salient nor very relevant to the task at hand. For instance, had we chosen to have health discussions about more gendered topics, then the coach's or client's gender would likely have had a higher salience compared to the generic health topics that were discussed (e.g., sleep quality, nutrition, and exercise, which affect everybody). Along similar lines, the current study included only a single interaction with the ECA, which naturally limits the degree to which gender-compatibility can affect the nascent relationship. For comparison, Schmalbach et al. [31] examined gender-matching in the context psychotherapy outcomes. Sessions took place over periods lasting from three months to six years. However, while these authors found significant gender-matching effects on treatment outcomes (quality of life and symptom reduction), the effects were weak, likely contingent on other factors, and only detectable with a much larger sample than ours. The current sample comprised 40 participants in a between subjects design and thus would likely require either a larger sample or additional efforts to boost the effect of gender matching. Going forward, we foresee that one could e.g. have more long-term interactions, or even interactions with alternating coaches.

Another clear limitation pertains the speed of turn-taking and other aspects that negatively impacted the interaction's naturalness. In particular, technical constraints regarding API-response-time and speech-to-text conversion introduced significant delays between the clients' speech and the system's response. Although we instructed participants in advance about this and they generally accepted it, it is still a limitation worth mentioning as it certainly prevented a smoother interactional flow. Going forward, there are several ways to address this: First, technology is swiftly advancing, thereby reducing delay times. Second, it would be possible to convert individual sentences to speech rather than waiting for the entire LLM-response. Third, there could be filler-activities, such as note-taking by the coach, which could give the pause some acceptable meaning.

Both limitations - regarding effect strength/sample size and interaction naturalness - are naturally springboards for future research: First, we see a large need for follow-up studies that have a higher frequency and intensity, focusing, for example, on more consequential topics (e.g., student stress consulting during finals week as opposed to a general health coaching session), and follow up with a longer-term and outcome-oriented horizon (e.g., connecting to grades, symptom reduction, etc., beyond the currently used ad-hoc evaluation metrics). Second, there is room to improve the non- and paraverbal behavior of the agent. In the current study, only basic nonverbals were included (lip-sync, general orientation-following). Given that nonverbal factors are key to the development of rapport, improving this aspect of the agent is critical.

6. Conclusion

This was a proof-of-concept study that showed the feasibility of designing effective LLM-based ECAs for health support and interventions. We conducted an experiment where we manipulated human-ECA similarity via gender-matching and examined three research questions. The quantitative evaluation showed that human-ECA similarity either had no effect or actually lowered their satisfaction with the interaction. The qualitative evaluation showed nuanced perspectives about the effectiveness of LLM-based ECAs. Overall, our study sets a foundation for future work on artificial social influence, or how intelligent agents influence human judgement and behavior.

References

[1] J. Cassell, Embodied conversational interface agents, Communications of the ACM 43 (2000) 70–78.

[2] B. Lugrin, Introduction to socially interactive agents, in: The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 1: Methods, Behavior, Cognition, Association for Computing Machinery, 2021, pp. 1–20.

[3] J. Russell Stuart, P. Norvig, Artificial intelligence: a modern approach, Prentice Hall Upper Saddle River, 2009.

[4] J. Cassell, K. R. Thorisson, The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents, Applied Artificial Intelligence 13 (1999) 519–538.

[5] N. Krämer, S. Kopp, C. Becker-Asano, N. Sommer, Smile and the world will smile with you—the effects of a virtual agent's smile on users' evaluation and behavior, International Journal of Human-Computer Studies 71 (2013) 335–349.

[6] T. W. Bickmore, R. A. Silliman, K. Nelson, D. M. Cheng, M. Winter, L. Henault, M. K. Paasche-Orlow, A randomized controlled trial of an automated exercise coach for older adults, Journal of the American Geriatrics Society 61 (2013) 1676–1683.

[7] P. Gardiner, A. S. Lestoquoy, K. Gergen-Barnett, B. Penti, L. F. White, R. Saper, L. Fredman, S. Stillman, N. L. Negash, P. Adelstein, et al., Design of the integrative medical group visits randomized control trial for underserved patients with chronic pain and depression, Contemporary clinical trials 54 (2017) 25–35.

[8] P. Philip, J.-A. Micoulaud-Franchi, P. Sagaspe, E. D. Sevin, J. Olive, S. Bioulac, A. Sauteraud, Virtual human as a new diagnostic tool, a proof of concept study in the field of major depressive disorders, Scientific reports 7 (2017) 42656.

[9] M. Y. Baihaqi, A. G. Contreras, S. Kawano, K. Yoshino, Rapport-driven virtual agent: Rapport building dialogue strategy for improving user experience at first meeting, arXiv preprint arXiv:2406.09839 (2024).

[10] J. W. Ayers, A. Poliak, M. Dredze, E. C. Leas, Z. Zhu, J. B. Kelley, D. J. Faix, A. M. Goodman, C. A. Longhurst, M. Hogarth, et al., Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum, JAMA internal medicine 183 (2023) 589–596.

[11] Z. Ma, Y. Mei, Z. Su, Understanding the benefits and challenges of using large language model-based conversational agents for mental well-being support, in: AMIA Annual Symposium Proceedings, volume 2023, American Medical Informatics Association, 2023, p. 1105.

[12] F. J. Bernieri, Coordinated movement and rapport in teacher-student interactions, Journal of Nonverbal behavior 12 (1988) 120–138.

[13] J. Gratch, G. Lucas, Rapport between humans and socially interactive agents, in: The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 1: Methods, Behavior, Cognition, Association for Computing Machinery (ACM), 2021, pp. 433–462.

[14] L. Tickle-Degnen, R. Rosenthal, The nature of rapport and its nonverbal correlates, Psychological inquiry 1 (1990) 285–293.

[15] F. Xie, A. Derakhshan, A conceptual review of positive teacher interpersonal communication behaviors in the instructional context, Frontiers in psychology 12 (2021) 708490.

[16] J. N. Cappella, On defining conversational coordination and rapport, Psychological Inquiry 1 (1990) 303–305.

[17] B. N. Frisby, M. M. Martin, Instructor–student and student–student rapport in the classroom, Communication Education 59 (2010) 146–164.

[18] C. M. Estepp, T. G. Roberts, Teacher immediacy and professor/student rapport as predictors of motivation and engagement, Nacta Journal 59 (2015) 155–163.

[19] J. A. Harrigan, T. E. Oxman, R. Rosenthal, Rapport expressed through nonverbal behavior, Journal of nonverbal behavior 9 (1985) 95–110.

[20] G. W. Joe, D. D. Simpson, D. F. Dansereau, G. A. Rowan-Szal, Relationships between counseling rapport and drug abuse treatment outcomes, Psychiatric services 52 (2001) 1223–1229.

[21] M. J. Leach, Rapport: A key to treatment success, Complementary therapies in clinical practice 11 (2005) 262–265.

[22] R. B. Cialdini, N. J. Goldstein, Social influence: Compliance and conformity, Annu. Rev. Psychol. 55 (2004) 591–621.

[23] J. Gratch, A. Okhmatovskaia, F. Lamothe, S. Marsella, M. Morales, R. J. van der Werf, L.-P. Morency, Virtual rapport, in: Intelligent Virtual Agents: 6th International Conference, IVA 2006,

Marina Del Rey, CA, USA, August 21-23, 2006. Proceedings 6, Springer, 2006, pp. 14–27.
[24] L. Huang, L.-P. Morency, J. Gratch, Virtual rapport 2.0, in: Intelligent Virtual Agents: 10th International Conference, IVA 2011, Reykjavik, Iceland, September 15-17, 2011. Proceedings 11, Springer, 2011, pp. 68–79.
[25] M. A. Hogg, D. J. Terry, K. M. White, A tale of two theories: A critical comparison of identity theory with social identity theory, Social psychology quarterly (1995) 255–269.
[26] M. McPherson, L. Smith-Lovin, J. M. Cook, Birds of a feather: Homophily in social networks, Annual review of sociology 27 (2001) 415–444.
[27] S. Lim, R. Schmälzle, G. Bente, Artificial social influence via human-embodied ai agent interaction in immersive virtual reality (vr): Effects of similarity-matching during health conversations, arXiv preprint arXiv:2406.05486 (2024).
[28] G. Bente, R. Schmälzle, N. T. Jahn, A. Schaaf, Measuring the effects of co-location on emotion perception in shared virtual environments: An ecological perspective, Frontiers in Virtual Reality 4 (2023) 1032510.
[29] A. M. Von der Pütten, N. C. Krämer, J. Gratch, S.-H. Kang, "it doesn't matter what you are!" explaining social effects of agents and avatars, Computers in Human Behavior 26 (2010) 1641–1650.
[30] V. Braun, V. Clarke, Thematic analysis., American Psychological Association, 2012.
[31] I. Schmalbach, C. Albani, K. Petrowski, E. Brähler, Client-therapist dyads and therapy outcome: Does sex matching matter? a cross-sectional study, BMC Psychology 10 (2022) 52. doi:10.1186/s40359-022-00769-3.

A. Appendix A

Participants received these instructions for the get-to-know-you and health consultation task.

Instruction for the Get-to-Know Task

For the first task, you will meet with a health coach, who is represented by an embodied agent, and talk freely just like how you would engage with a person you meet for the firat time. The task will last for 5 minutes.

Please wait until the agent speaks, and then start talking once it finishes. Once the five minutes is complete, the screen will freeze and a rotating hour-glass will appear on the screen.

The agent will always do their best to think of effective responses to you. That means it will take about 5-10 seconds for the agent to process and respond to you. Below are some guidelines to make the conversation flow well:

1. Say everything you want to say in one sentence without pauses.
2. While the agent is processing what you said, take the time to think of what you want to talk about with the agent next. This will help keep the conversation going.
3. Sometimes the agent may stop talking mid-sentence. If this happens, you can either start a new topic/ask a new question, or ask the agent to finish what it was saying.
4. Please start speaking once the agent finishes speaking. Since the agent is so excited to talk with you, it may not register that you have interrupted them.

Finally: Your privacy is important. Any personal information shared will be kept confidential and used for research purposes only.

Thank you for your participation!

Figure 2: Instructions for the Get-to-Know-You Task.

Instruction for the Consultation Task

For the next task, you will meet with a health coach, who is represented by an embodied agent, for a free consultation about nutrition. The task will last for 5 minutes.

Your task is to discuss your nutritional habits, preferences, and any specific goals you have in this area. You can simply engage in a conversation about your typical daily meals and snacks, ask the coach questions about healthy diets, or explore areas where you think there is room for improvement in how and what you eat.

Again, please wait until the agent speaks, and then start talking once it finishes. Once the five minutes is complete, the screen will freeze and a rotating hour-glass will appear on the screen.

As a reminder, the agent will always do their best to think of effective responses to you. That means it will take about 5-10 seconds for the agent to process and respond to you. Below are some guidelines to make the conversation flow well:

1. Say everything you want to say in one sentence without pauses.
2. While the agent is processing what you said, take the time to think of what you want to talk about with the agent next. This will help keep the conversation going.
3. Sometimes the agent may stop talking mid-sentence. If this happens, you can either start a new topic/ask a new question, or ask the agent to finish what it was saying.
4. Please start speaking once the agent finishes speaking. Since the agent is so excited to talk with you, it may not register that you have interrupted them.

Finally: Your privacy is important. Any personal information shared will be kept confidential and used for research purposes only.

Thank you for your participation!

Figure 3: Instructions for the Consultation Task.