Emotional Responses to Exclusionary Behaviors in Intelligent Embodied Augmented Reality Agents

K. Apostolou[1], V. Milata[1], F. Škola[1], and F. Liarokapis[1]

[1]CYENS Centre of Excellence, Cyprus

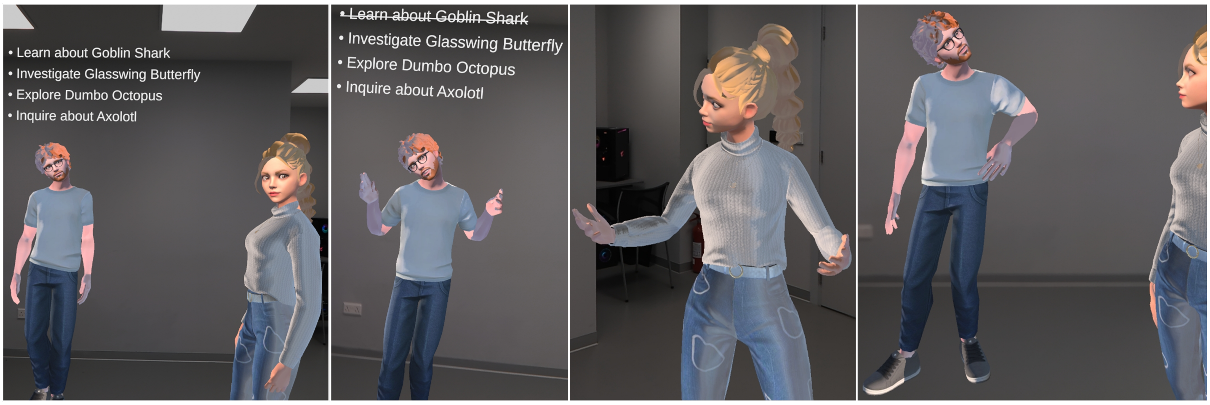

Figure 1: *Illustration of the interaction between two AI agents and the user: The first frame introduces the* responder *(male) and* inquirer *(female) agents' avatars. In the second, the* responder *addresses the user's initial objective. The third features the* inquirer *posing a question for a distinct set of objectives, and the fourth showcases communication between the* responder *and* inquirer.

Abstract

This study investigated how interactions with intelligent agents, embodied as augmented reality (AR) avatars displaying exclusionary behaviors, affect users' emotions. Six participants engaged using voice interaction in a knowledge acquisition scenario in an AR environment with two ChatGPT-driven agents. The gaze-aware avatars, simulating realistic body language, progressively demonstrated social exclusion behaviors. Although not statistically significant, our data suggest a post-interaction emotional shift, manifested by decreased positive and negative affect–aligning with previous studies on social exclusion. Qualitative feedback revealed that some users attributed the exclusionary behavior of avatars to system glitches, leading to their disengagement. Our findings highlight challenges and opportunities for embodied intelligent agents, underscoring their potential to shape user experiences within AR, and the broader extended reality (XR) landscape.

CCS Concepts

• ***Human-centered computing*** → ***Empirical studies in collaborative and social computing;*** • ***Applied computing*** → *Psychology;* • ***Theory of computation*** → *Models of learning;*

1. Introduction

The rapid growth in Artificial Intelligence (AI) has led to the development of advanced Large Language Models (LLMs), such as the Generative Pre-trained Transformer (GPT). These models, operating on generative AI principles, use deep learning and neural networks to find patterns in existing data. The versatility of their outputs mirrors human creativity, highlighting not only their potential to revolutionize how machines comprehend and generate natural language [Ray23] but also their ability to understand and mimic human emotions [EHSAL23, LJGZM*24].

This study leverages the power of the LLMs (ChatGPT with the GPT-3.5 model) to drive the behavior of embodied agents in

DOI: 10.2312/egs.20241023

an Augmented Reality (AR) environment. Realistic gaze-aware avatars capable of simulating body language were employed to act as GPT-driven agents in AR. An exploration of how these agents impact user experiences through simulated social behaviors was conducted, to gain insights into the changes of users' emotional states following interactions. Particularly, the focus of this study was to examine users' responses when intelligent agents in the AR space displayed social exclusion, behavior known to elicit numerous negative outcomes (reduced social participation [MBF*07, TBD*07], increased aggression [TBTS01], diminished cognitive performance [BTN02], and dehumanization [BH10] among others).

Our investigation was based on a question-answering knowledge acquisition scenario in AR, where participants engaged in spoken conversation with an AI agent designed to provide information (*responder*). At the same time, another AI agent tailored to seek information about its own set of objectives (*inquirer*; see Figure 1), was present in the AR environment. The interactions revolved around unfamiliar topics, tasking participants with addressing a predetermined list of objectives during the conversation. Notably, as the conversation concluded, the *responder*'s avatar exhibited exclusionary behavior by avoiding eye contact, disregarding user prompts, and solely interacting with *inquirer*.

Building on this scenario, the primary objective of the study was to evaluate the shift in users' emotional states (both positive and negative affect) following exclusionary behaviors and to qualitatively assess the potential implications of intelligent AR agents in shaping overall user experiences. Additionally, users' learning outcomes related to the provided objectives were evaluated by a test for assessment. This comprehensive examination dives into the emotional impact on users resulting from interactions with intelligent agents in AR, providing insights into the evolving field of human-AI interactions in immersive extended reality (XR) settings (augmented, virtual, and mixed reality).

2. Related Works

How people perceive embodied agents in XR tends to have similarities to how real people are perceived [MKA*19], and perception of virtual agents and human avatars do not seem to evoke different levels of social presence within XR [KGK*17]. When performing tasks in an AR environment that included embodied agents, participants showed changed levels of social connectedness, nonverbal communication, and task performance (following the pattern of social facilitation and inhibition) [MJH*19]. Moreover, mere proximity to embodied agents in AR leads to behavior similar to being close to a real human [HKC*22]. In terms of the relationship towards embodied intelligent agents in AR, research has shown that users have heightened confidence in them, compared to their non-AR (non-embodied) counterparts. This effect likely originates in the social component of the interaction, which was effectively mediated by locomotion, gesturing, and other human-like capabilities of the agents [KBH*18].

Social exclusion, while not a clearly defined phenomenon, is typically characterized by reduced participation in a social setting [MBF*07]. It has been demonstrated that social exclusion has detrimental effects on an individual's social behavior [TBD*07], leads to increased aggression [TBTS01], and self-defeating behavior [TCB02]. Beyond the social effects of this social phenomenon, a general decline in cognitive performance has been linked to social exclusion [BTN02], and it might even lead to dehumanization [BH10].

In a large meta-analysis of 192 studies on social exclusion, Blackhart et al. (2009) [BNKB09] found a notable shift towards a more negative emotional state caused by social exclusion. However, in terms of positive and negative affect, it was manifested by low levels of both Positive Affect (PA) and Negative Affect (NA). We hypothesize to find similar outcomes in the case of social exclusion mediated by embodied intelligent agents in the AR environment.

3. System Architecture

The application's front-end was implemented in Unity game engine (version 2022.3.1f1). *Responder* and *inquirer* agents were represented by Ready Player Me (`readyplayer.me`) avatars enhanced with eye-gazing and lip-syncing capabilities and also with various gesturing animations. Speech-to-Text and Text-to-Speech functionalities were implemented using Azure Cognitive Services. The Text-to-Speech service also generated visemes that were used to drive avatar lip-syncing. For AI capabilities driving natural language understanding and generation, ChatGPT with gpt-3.5-turbo-0613 (June 13th, 2023) model was used. Conversations between *inquirer* and *responder* were implemented using the same ChatGPT API calls as between the user and *responder*, to keep the communication dynamics (delays) consistent. The application was deployed on Magic Leap 2 AR head-mounted display. For integration of the key components, see Figure 2.

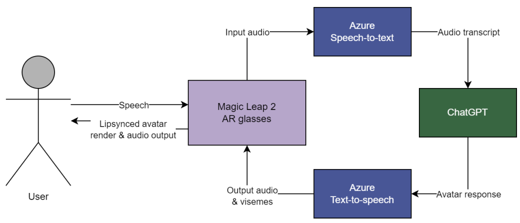

Figure 2: *Overview of the system architecture.*

4. Methodology

4.1. Participants

Six volunteers (five females) aged between 23 and 36 (M = 26.8, SD = 5.0) participated in a within-subjects design study. Three participants were familiar with AR, while all were familiar with virtual reality (VR). AR familiarity varied, with two participants having used it occasionally and one participant being a monthly user. Participants provided informed consent, but the true aim was concealed; they were informed they would interact with AI agents to acquire knowledge about specific objectives.

4.2. Procedure

The experimental AR environment contained two avatars embodying the AI agents; the *responder* (providing information) and the *inquirer* (seeking information about its own set of topics); see Figure 1). Participants were tasked to find information about selected objectives (exotic animals; glasswing butterfly, dumbo octopus, goblin shark, and axolotls) by interacting with the *responder* using natural language. The deliberate selection of these topics aimed to ensure participants had limited prior knowledge, enabling a thorough assessment of their learning outcomes. Participants were informed that the other agent (*inquirer*) would research its own set of objectives (also using natural language interaction with the *responder*), while the participant and *inquirer* would take turns interacting with the *responder*. However, the *responder* agent was designed to address the user's objectives initially, but eventually, it displayed exclusionary behaviors; i.e., responding to the *inquirer*'s objectives while ignoring the user's prompts, avoiding eye contact, and exhibiting negative body language.

Following human-AI interactions, participants were tested to evaluate learning effects. The test included six questions, including 4 multiple-choice questions such as "At what depths are Goblin Sharks found?" (ChatGPT system prompt was customized to specifically include the details that were present in the evaluation when a relevant question was asked). PA and NA scores were assessed using the PANAS (Positive and Negative Affect Schedule) questionnaire [WCT88] both before and after the experiment. Both PA and NA were assessed by summing the answers, resulting in a value between 10 and 50 for each category. Finally, participants were interviewed to assess their distress levels and perceptions of the interactions with the AR agents, specifically exploring feelings of exclusion.

4.3. Exclusionary Scenario

Initially, the *responder* was set up to turn to the participant and make eye contact when the question from the participant was expected. The first turn in asking had the participant, the *inquirer* asked its question after the first participant's question was answered. This was followed by the *responder* explicitly prompting the participant to ask more questions in the second round of asking-answering. In the third round, the *responder* did not prompt the user but made eye contact after *inquirer*'s question was answered (and the participant's question was answered as well).

The exclusionary behavior started in the fourth round of expected participant-*responder* interaction. The *responder* did not turn to or make eye contact with the participant and kept looking at the *inquirer*. If participants asked their question at this point, the *responder* would turn towards the user for 5 seconds, but the question was ignored. Instead, *responder* answered *inquirer*'s question that was asked at the end of the 5-sec idle period.

In the last round, *responder* still looks toward the *inquirer*. If the participant asked a question, *responder* would turn toward the user, but the question was ignored. After 5 sec, the *responder* turned toward the *inquirer* prompting it to ask more questions. *Inquirer* indicated it had no more questions, and *responder* in turn thanked for the interactions and the experiment ended.

5. Results

Median affect scores were 36 and 22 for PA and NA before the experiment, respectively. Post-experiment, participants reported median scores of 29 for PA and 17 for NA. This observed decrease in both PA and NA suggests a shift in participants' emotional states.

To examine the significance of changes in PA and NA scores, Wilcoxon signed-rank tests were conducted. The test for PA revealed no significant difference in scores before and after the experimental procedure (V = 16, p = 0.29). Similarly, the test for NA indicated no significant shift in scores (V = 16, p = 0.31). Thus, there is insufficient evidence to support a substantial change in emotional states between the two time points (pre and post-experimental procedure) as proposed in our hypothesis. These findings collectively suggest that the experimental procedure did not result in a significant alteration in participants' PA or NA. See Figure 3.

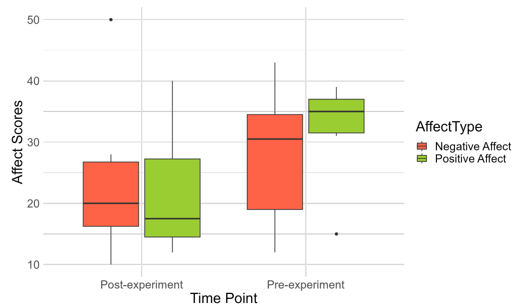

Figure 3: *Comparison of PANAS state positive/negative affect (PA and NA) distribution at baseline (pre) and after the experimental manipulation (post).*

An assessment of test outcomes revealed a moderate level of learning among participants, with a median performance of 66.66% (with a range from 50% to 100%). Qualitative data from participants' responses indicated that exclusionary interactions with AI agents led to perceived system faults and subsequent disengagement. Most participants attributed the agent's ignoring behavior to technical issues. Regarding learning effects, participants reported a predominant focus on avatar interactions over the informational context, potentially contributing to underperformance in the test.

6. Discussion

The findings of this study, offering both quantitative and qualitative perspectives on participants' experiences during human-AI interactions, bring attention to the observed decrease in both PA and NA scores, despite the absence of statistical significance. This shift in affect is in line with the previous studies on social exclusion outside the XR setting, with real people, suggesting a similar outcome in the case of the AR set-up with intelligent agents. Interestingly, despite participants reporting the negative behaviors as glitches, the mere exposure to what they perceived as exclusionary actions might have triggered a subtle yet meaningful shift in their emotional states, manifested by the decrease in both PA and NA scores. The experience of exclusion, even if interpreted as a technical malfunction, may have subtly influenced their self-perception within the context of the interaction.

Proceedings published by Eurographics - The European Association for Computer Graphics.

Furthermore, participants' interpretation of negative behaviors as glitches introduces a dimension related to trust in technology. The initial inclination to perceive issues as technical glitches rather than intentional actions could indicate a baseline level of trust in the system. Repeated experiences of perceived glitches could erode the trust participants have in the AR agents, emphasizing the need for sustained reliability of the system. A similar trend is visible from the perspective of knowledge acquisition, where the moderate level of learning demonstrated by participants suggests that participants' perceptions of AI agents' behavior (experiences of exclusionary interactions, coupled with perceived system faults), led to disengagement. This underscores the crucial need to address issues related to AI responsiveness and behavior, as these factors significantly shape user experiences.

While a low sample size limits the present study, it hints toward future directions, which should focus on disentangling the effect of diminished trust in technology from the effects of experimentally induced social exclusion. Differentiating these effects will enable a more detailed analysis of the intricate interplay between social psychology and human-computer interaction. Improving the experimental design to avoid perceptions of system glitches will facilitate the study of exclusionary behaviors, ostracism, and other problematic social patterns. Additionally, exploring or exploiting the similarity in the perception of user avatars and embodied nonhuman agents in XR environments would be intriguing in the context of the investigated topic.

7. Conclusion

This study examined the effects of exclusionary behavior demonstrated by intelligent AI agents embodied as AR avatars. Our results underscore the complex interplay between social psychology, trust dynamics, and learning outcomes in human-computer interactions. Despite the absence of statistical significance, the observed shift in emotional states, reflected in decreased PA and NA scores, implies effects akin to social exclusion observed outside XR environments, with real people. Participants exhibited limited levels of learning from the interaction, emphasizing the necessity for empathetic embodied agents. The conclusions of this study are limited by the low sample size; however, the presented results can be exploited as a stepping stone for more focused studies examining the impact of "misbehaving" embodied intelligent agents in the AR, and possibly the broader XR space.

8. Acknowledgements

This work has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 739578 and the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation and Digital Policy. It has also received funding from the European Union under grant agreement No 101093159. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

References

[BH10] Bastian B., Haslam N.: Excluded from humanity: The dehumanizing effects of social ostracism. *Journal of experimental social psychology 46*, 1 (2010), 107–113. 2

[BNKB09] Blackhart G. C., Nelson B. C., Knowles M. L., Baumeister R. F.: Rejection elicits emotional reactions but neither causes immediate distress nor lowers self-esteem: A meta-analytic review of 192 studies on social exclusion. *Personality and Social Psychology Review 13*, 4 (2009), 269–309. 2

[BTN02] Baumeister R. F., Twenge J. M., Nuss C. K.: Effects of social exclusion on cognitive processes: anticipated aloneness reduces intelligent thought. *Journal of personality and social psychology 83*, 4 (2002), 817. 2

[EHSAL23] Elyoseph Z., Hadar-Shoval D., Asraf K., Lvovsky M.: Chatgpt outperforms humans in emotional awareness evaluations. *Frontiers in Psychology 14* (2023). 1

[HKC*22] Huang A., Knierim P., Chiossi F., Chuang L. L., Welsch R.: Proxemics for human-agent interaction in augmented reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (2022), pp. 1–13. 2

[KBH*18] Kim K., Boelling L., Haesler S., Bailenson J., Bruder G., Welch G. F.: 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)* (2018), IEEE, pp. 105–114. 2

[KGK*17] Kothgassner O. D., Griesinger M., Kettner K., Wayan K., Völkl-Kernstock S., Hlavacs H., Beutl L., Felnhofer A.: Real-life prosocial behavior decreases after being socially excluded by avatars, not agents. *Computers in Human Behavior 70* (2017), 261–269. 2

[LJGZM*24] Llanes-Jurado J., Gómez-Zaragozá L., Minissi M. E., Alcañiz M., Marín-Morales J.: Developing conversational virtual humans for social emotion elicitation based on large language models. *Expert Systems with Applications 246* (2024), 123261. 1

[MBF*07] Morgan C., Burns T., Fitzpatrick R., Pinfold V., Priebe S.: Social exclusion and mental health: conceptual and methodological review. *The British Journal of Psychiatry 191*, 6 (2007), 477–483. 2

[MJH*19] Miller M. R., Jun H., Herrera F., Yu Villa J., Welch G., Bailenson J. N.: Social interaction in augmented reality. *PloS one 14*, 5 (2019), e0216290. 2

[MKA*19] Mousas C., Koilias A., Anastasiou D., Rekabdar B., Anagnostopoulos C.-N.: Effects of self-avatar and gaze on avoidance movement behavior. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (2019), IEEE, pp. 726–734. 2

[Ray23] Ray P. P.: Chatgpt: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems 3* (2023), 121–154. 1

[TBD*07] Twenge J. M., Baumeister R. F., DeWall C. N., Ciarocco N. J., Bartels J. M.: Social exclusion decreases prosocial behavior. *Journal of personality and social psychology 92*, 1 (2007), 56. 2

[TBTS01] Twenge J. M., Baumeister R. F., Tice D. M., Stucke T. S.: If you can't join them, beat them: effects of social exclusion on aggressive behavior. *Journal of personality and social psychology 81*, 6 (2001), 1058. 2

[TCB02] Twenge J. M., Catanese K. R., Baumeister R. F.: Social exclusion causes self-defeating behavior. *Journal of personality and social psychology 83*, 3 (2002), 606. 2

[WCT88] Watson D., Clark L. A., Tellegen A.: Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology 54*, 6 (1988), 1063–1070. 3