International Journal of Human–Computer Interaction

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hihc20

Exploring the Effects of Mixed Reality Agents' Locomotion and Postures on Social Perception Through a Board Game

Zhuang Chang, Jiashuo Cao, Kunal Gupta, Huidong Bai & Mark Billinghurst

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2024.2435694

Taylor & Francis
Taylor & Francis Group

Exploring the Effects of Mixed Reality Agents' Locomotion and Postures on Social Perception Through a Board Game

Zhuang Chang[a], Jiashuo Cao[a], Kunal Gupta[a], Huidong Bai[a], and Mark Billinghurst[a,b]

[a]Auckland Bioengineering Institute, The University of Auckland, Auckland, New Zealand; [b]STEM, The University of South Australia, Adelaide, Australia

ABSTRACT

Non-verbal cues like locomotion and posture influence users' perceptions of Mixed Reality Agents (MiRAs). While Electroencephalography (EEG) captures cognitive responses, the influence of MiRAs' locomotion and postures on brain activity remains underexplored. Additionally, few studies integrate subjective and behavioral measures with EEG to evaluate these cues' impact on social perception. To address this, we conducted a within-subject study where participants played Gobang against three virtual agents in mixed reality: 1) a speech-only agent (**S**), 2) an embodied agent with speech and locomotion (**S + L**), and 3) an embodied agent with speech, locomotion, and posture (**S + L + P**). Results showed the **S + L + P** agent had higher engagement measured by the questionnaire but a lower EEG-based engagement index at AF3 than the **S + L** agent. Besides, the **S + L + P** was also rated higher in social presence, engagement, and emotional arousal than the **S** condition; No behavioral differences were observed. We discuss how MiRAs' locomotion and posture affect users' social perception and provide design implications for future human-agent interactions.

KEYWORDS

Virtual agents; EEG; engagement; mixed reality

1. Introduction

Integrating Intelligent Virtual Agents (IVAs) into applications has significantly transformed the landscape of Human-Computer Interaction (HCI). IVAs are typically classified into two types: 1) Voice-only agents (Katsarou et al., 2023) with no embodiment, like Amazon Alexa,[1] Microsoft Cortana,[2] and Apple Siri,[3] and 2) embodied agents (Yousefi et al., 2024) which possess a body, like Soul Machines' conversational virtual characters[4] and Non-Player Characters (NPCs) in computer games (Pretty et al., 2024). Despite the commercial success of voice-only agents, previous research has demonstrated that embodied agents can further enhance human-agent social interactions (Kruse et al., 2023; Mendes et al., 2024).

Therefore, we focus primarily on embodied agents in this work. IVAs embodied in the Mixed Reality (MR) environment are called Mixed Reality Agents (MiRAs) (Holz et al., 2011), which inhabit users' real-world environments. Unlike voice-only agents that communicate with users primarily through speech, MiRAs can further utilize non-verbal cues, such as facial expression, posture, and locomotion, to convey more complex social signals like emotions and behavioral intentions (Norouzi et al., 2020; Wang & Ruiz, 2021).

Specifically, locomotion and posture are two important non-verbal cues in MiRAs that could influence users' social perceptions. For example, a virtual agent walking within the user's immediate environment enhances the naturalness and intuitiveness of its behavior, thereby increasing its overall believability (Kim et al., 2018a). Additionally, postures presented by MiRAs have been shown to affect users' willingness to interact with these agents (Li et al., 2018).

Despite these findings, previous studies have predominantly relied on subjective and behavioral measures to assess the impact of MiRAs' non-verbal cues on social perception. Since all cognitive activities are driven by the brain, physiological measures like Electroencephalography (EEG), which records electrical brain activity, have been used to assess cognitive load (Kumar & Kumar, 2016), trust in virtual assistants (Gupta et al., 2020), and engagement in gaming (Pope et al., 1995; Ruqeyya et al., 2022). In the human-agent interaction domain, EEG has also been employed to monitor real-time user engagement, allowing robots to adapt their behaviors to recapture users' diminishing attention levels (Szafir & Mutlu, 2012; Vrins et al., 2022). However, it remains unclear how virtual agents' non-verbal cues, such as posture and locomotion in mixed reality spaces, influence user engagement as measured by EEG.

In this work, we conduct a within-subjects user study to explore how the IVAs' locomotion and postures influence the user's brain activity and feelings of social presence and engagement in MR. Inspired by the iCat affective chess robot, which engages users in a turn-based chess game where the agent strategizes and plays against users (Leite et al., 2009), we developed a turn-based Gobang game in MR. Gobang was chosen because it provides a simple yet strategic framework for evaluating user interactions with

CONTACT Zhuang Chang zcha621@aucklanduni.ac.nz; Mark Billinghurst mark.billinghurst@auckland.ac.nz Auckland Bioengineering Institute, The University of Auckland, 70 Symonds Street, Auckland, New Zealand

virtual agents, allowing us to investigate the impact of non-verbal cues, such as locomotion and posture, in a dynamic, competitive setting. Similar to Kim et al. (2018a), which used a speech-only agent as a comparison condition to demonstrate the benefits of MiRAs, we studied three corresponding types of virtual agents: 1) **S**peech-only agent (**S**), 2) Embodied agent with **S**peech and **L**ocomotion (**S + L**), and 3) Embodied agent with **S**peech, **L**ocomotion, and **P**ostures (**S + L + P**). Our results indicate that locomotion and postures enhanced social presence and social engagement but reduced the EEG-measured engagement index. Our work contributes in three ways:

- C1. We explore the effects of MiRAs' non-verbal cues on user engagement by integrating EEG with subjective and behavioral measurements;
- C2. We present a user study to examine the influence of MiRAs' locomotion and postures on users' social perceptions;
- C3. We provide novel design implications for human-agent interaction in MR, derived from our comprehensive analysis of experimental results and observations.

The remainder of this paper is organized as follows: Section 2 reviews related work and highlights the novelty of our research. Section 3 introduces the system used for the experiment, followed by the user study design in Section 4. In Section 5, we present the experimental results, while Section 6 provides a discussion of the findings. Section 7 outlines the limitations of the study, and finally, Section 8 offers conclusions and suggests directions for future work.

2. Related work

In this study, we investigate how MiRAs' locomotion and postures affect social perception within the context of a board game. To establish the foundation for our research, we review related work in three key areas: 1) Non-verbal cues in embodied virtual agents, 2) virtual agents in board games, and 2) EEG-based engagement index.

2.1. Non-verbal cues in embodied virtual agents

Non-verbal cues in embodied virtual agents, such as posture (Li et al., 2018), eye gaze (Kevin et al., 2018), facial expressions (Milcent et al., 2022) and proxemics (Ye et al., 2021), can be used to enhance humanagent social interaction (Miller et al., 2019; Wang & Ruiz, 2021). For example, DeVault et al. (2014) designed a virtual therapist for an engaging face-to-face interview scenario where users felt comfortable talking and sharing information. To establish the relationship between the virtual therapist and the user, they considered facial expressions, gaze, gestures, and postures when designing the virtual agent.

Among the various non-verbal cues, posture is one of the most influential in embodied virtual agents. It can shape the agent's personality (Antonio Gómez Jáuregui et al., 2021; Ishii et al., 2020), enhance users' sense of co-presence (Chang et al., 2022), and increase engagement (Kum & Lee, 2022). For instance, Li et al. (2018) explored the effects of agent posture and embodiment on social distance in MR, finding that users maintained greater social distance from agents displaying open postures, yet were more willing to interact with them. Similarly, Randhavane et al. (2019) developed an algorithm to model perceived friendliness in Augmented Reality (AR) by adjusting an agent's gaze, gait, and gestures such as head nodding and waving. This significantly enhanced both friendliness and social presence. Chang et al. (2022) also found that embodied agents in immersive Virtual Reality (VR) with posture cues increased co-presence and captured more attention, as evidenced by EEG measurements. These studies demonstrate that the postures of virtual agents in 3D immersive environments (VR/AR/MR) can significantly influence users' social perceptions in various aspects, including social distance, perceived friendliness, social presence, co-presence, and attention. More relevant studies could be found in (Kruse et al., 2023; Norouzi et al., 2020; Yousefi et al., 2024).

In addition to posture, locomotion is a critical component of human-agent interaction, especially for virtual agents embodied in 3D immersive AR, VR, and MR environments (Holz et al., 2011; Nijholt, 2022; Norouzi et al., 2019). Like posture, locomotion can also influence users' social perceptions. For instance, Kim et al. (2018a) found that a virtual agent walking in an AR space elicited higher social presence, engagement, and trust compared to a speech-only agent. Locomotion, which involves changes in spatial positioning, also affects proxemics, the physical distance users maintain from virtual agents during interactions (Huang et al., 2022; Lee et al., 2018; Miller et al., 2019). Moreover, locomotion can convey a virtual agent's awareness of the virtual and physical environments, further enhancing its believability and improving user experiences. For example, Ye et al. (2021) introduced a real-time position-aware locomotion method for room-scale VR, while Kim et al. (2017) demonstrated that a virtual human in AR avoiding physical obstacles like a chair increased social presence and influenced users' behaviors, leading them to avoid walking through the virtual human's space. In short, locomotion in embodied virtual agents affects users' social perceptions and behaviors.

Although previous research has investigated the effects of postures and locomotion in MiRAs, they primarily employed subjective and behavioral measurements. In our work, we evaluate the effects of MiRAs' locomotion and postures by integrating subjective, behavioral, and physiological measurements (i.e., EEG). We use locomotion to indicate the virtual agent's intended spatial positions in the MR space and express the virtual agent's emotional state through postures. With the designed locomotion and postures, we mainly focus on manipulating user engagement in the Gobang game and measuring the engagement using both questionnaires and EEG.

2.2. Virtual agents in board games

Board games, typically multi-player and turn-based, have been widely used as environments for studying human-agent interactions (Barambones et al., 2023; Damette et al., 2024; Piette et al., 2021). For instance, Sun et al. (2022) implemented the Colored Trails board game in VR to investigate how a virtual robotic opponent's theory of mind abilities influenced users' delegation behavior, requiring participants to negotiate with the agent to succeed. Such a structured, rule-based environment provides an ideal testbed for exploring various aspects of human behavior in interaction with virtual agents.

In board games, virtual agents can serve either as opponents, competing against users, or as assistants, helping users win the game. For instance, Eichhorn et al. (2021) designed virtual chess opponents that train players in strategies on a tablet screen, while Allameh and Zaman (2021) developed a virtual assistant on screen for The Royal Game of Ur, that could answer rule-related questions, provide strategic move suggestions based on the player's state, and notify players of important game events. Because of the clear rules and structured nature of board games, these virtual agents can effectively monitor game flow, enhancing user engagement.

In contrast to screen-based virtual agents, those embodied in VR, AR, or MR environments have been integrated into board games, providing a more immersive and interactive experience (Lee et al., 2021; Liu et al., 2023; Torre et al., 2000). For example, Lee et al. (2018) introduced a virtual human in AR that sits at a physical table to play a tabletop game, where each player takes turns moving their tokens along designated spots on the shared surface. This setup enhances the sense of presence and facilitates more natural interactions between players and the virtual agent.

Despite advancements in integrating AR/MR virtual agents into board games, the limited field of view (FoV) of headsets has frequently been reported to impair user experience, as users cannot view the virtual agent's entire body while focusing on the game board (Lee et al., 2018; Wang et al., 2019). One potential solution is to employ miniature virtual agent bodies (Kim et al., 2020; Wang et al., 2019). Alternatively, scaling up the board to a giant size placed on the ground allows players to walk on top, thereby enhancing visibility and interaction (Bocchi et al., 2024; Mouton et al., 2017).

In our study, we integrated miniature virtual opponents into a scaled-up Gobang board game (Li et al., 2022) in the MR environment, enabling agents to navigate and place pieces on the enlarged game board. The rationale for selecting Gobang as a testbed is further discussed at the beginning of section 3 and subsection 3.2.

2.3. EEG-based engagement index

To measure user engagement, we recorded their brain activity using EEG. EEG has previously been used to measure cognitive load (Antonenko et al., 2010), visual attention (Busch & VanRullen, 2010), and engagement (McMahan et al., 2015). Pope et al. (Pope et al., 1995) developed an EEGbased engagement index based on the ratio of beta power (13-22 Hz) and the sum of alpha power (8-13 Hz) and theta power (4-8 Hz). They applied it in a closed-loop system to adjust the user task allocation to maintain higher task engagement. Similarly, a further study (Freeman et al., 1999) extended this work in two experiments with larger sample sizes and more detailed observations. Their results confirmed the effectiveness of this EEG-based engagement index.

The EEG-based engagement index has also been widely used to measure human cognitive states in later research (Berka et al., 2007; McMahan et al., 2015; Nuamah & Seong, 2018; Szafir & Mutlu, 2012). For example, Berka et al. (2007) measured engagement and workload using EEG during different types of standard cognitive tests. They argued that the EEG engagement index reflects information-gathering, visual processing, and attention allocation, while the EEG workload is sensitive to working memory load, integration of information, and analytical reasoning. Szafir and Mutlu (2012) designed an adaptive agent that could monitor student attention in real time using the EEG-based engagement index. Similarly, McMahan et al. (2015) used the EEG-based engagement index to measure players' cognitive engagement during video games. They found that the engagement index was effective in differentiating high-intensity game events from general gameplay. Moreover, Nuamah and Seong (2018) identified significant differences in task engagement indices derived from EEG signals recorded from participants performing five distinct cognitive tasks. In conclusion, the EEG-based engagement index has been proven to be effective in measuring cognitive engagement.

Although there are other ways to measure engagement based on EEG signals (Coelli et al., 2015), in this paper, we used the EEG-based engagement index proposed by Pope et al. (1995) to observe participants' cognitive engagement as it has been widely used in later research (Castiblanco Jimenez et al., 2022; Dehais et al., 2020; Rajendran et al., 2022).

In summary, locomotion and postures are crucial non-verbal cues in MiRAs, yet prior research has largely relied on subjective and behavioral measures to assess their impact. Board games provide a structured and rule-based environment that is ideal for studying human interactions with virtual agents. EEG effectively captures users' cognitive activities and engagement levels, offering valuable physiological insights. In this study, we integrate subjective, behavioral, and physiological (EEG) measurements to explore how MiRAs' locomotion and postures influence users' social perceptions within a Gobang game environment.

3. System overview

Our system primarily consists of the IVAs and the MR Gobang game, where the IVAs compete against users in the MR Gobang game on the HoloLens 2. Previous studies have noted that the limited field of view (FoV) of the HoloLens 2 impacts users' experiences with virtual agents (Lee et al., 2021; Li et al., 2018), and that miniature embodied agents walking around can appear more approachable and relatable to users (Wang et al., 2019). Besides, the limited FoV may lead to more user head movements, which could cause more noise in the EEG signals. Therefore, in our system, we designed the embodied agent as a miniature character capable of navigating the game board and placing chess pieces, emulating a player's actions in a giant chess game (Bocchi et al., 2024).

In the following subsections, we will describe the IVAs, the MR Gobang game, and the system implementation.

3.1. The intelligent virtual agents

In our study we used three difference IVAs: 1) A **S**peech-only agent (**S**), 2) an embodied agent with **S**peech and **L**ocomotion (**S + L**), and 3) an embodied agent with **S**peech, **L**ocomotion and **P**ostures (**S + L + P**).

IVA **S** (**S**peech): The speech-only agent was designed to interact with people through voice only, similar to the Speech agent used by Kim et al. (2018a), we also pre-generated the speech using the IBM Watson Text-to-Speech service[5] before interacting with the virtual agent to avoid any network delay during the real-time text to speech generation (Chang et al., 2022). To increase the speech richness, we prepared synonymous sentences for the same speech intention and randomly picked one when a specific event requiring speech intention happened. For example, when the virtual agent sees the user (event), it would greet the user (speech intention) by randomly picking one speech sample from "Hi, how are you", "Hello, nice to meet you" or "Hi, what's up". We used a unique keyword to describe the speech intention (e.g., "greeting") and mapped the keyword with a value list consisting of several synonymous sentences. All of the pre-generated audio clips were named with the combination of a unique keyword and an index of a sentence in the synonymous sentence list. For example, the names "greeting1", "greeting2", and "greeting3" could represent three greeting audio clips generated with three synonymous sentences.

IVA **S + L** (**S**peech and **L**ocomotion): The speech and locomotion agent was designed based on the IVA speech condition, but included an embodied agent. The virtual agent had a full virtual human body and could walk toward target points with walking animations. By default, the virtual human kept facing the users with a slight body movement to avoid an impression of stiffness. When the virtual agent starts walking, it turns the body to face the target place and then walks towards the target. Once the virtual agent arrives at the target point, it turns around to face the user. The virtual character's walking behavior primarily utilizes a Mixamo[6] -generated walking animation and involves positional adjustments over time using the DoTween[7] Unity animation plugin. These adjustments are computed by dividing the distance between the character's starting and target positions by a constant speed value, ensuring smooth movement at a consistent pace. The virtual agent could also gaze at the user to maintain mutual gaze while being looked at, which was designed to enhance the realism of the interaction (Gregory et al., 2021). We used the same method as shown in Chang et al. (2022) to generate the virtual character model. We also used the Lipsync visemes[8] to synchronize the lip movement with the speech sample.

IVA **S + L+P** (**S**peech, **L**ocomotion and **P**osture): The speech, locomotion, and posture agent was based on the IVA S + L, but also using postures to express anger, happiness, sadness, thinking behaviors and a neutral expression as is illustrated in the Figure 1. Like the walking animation, all of these posture animations were also downloaded from Mixamo.

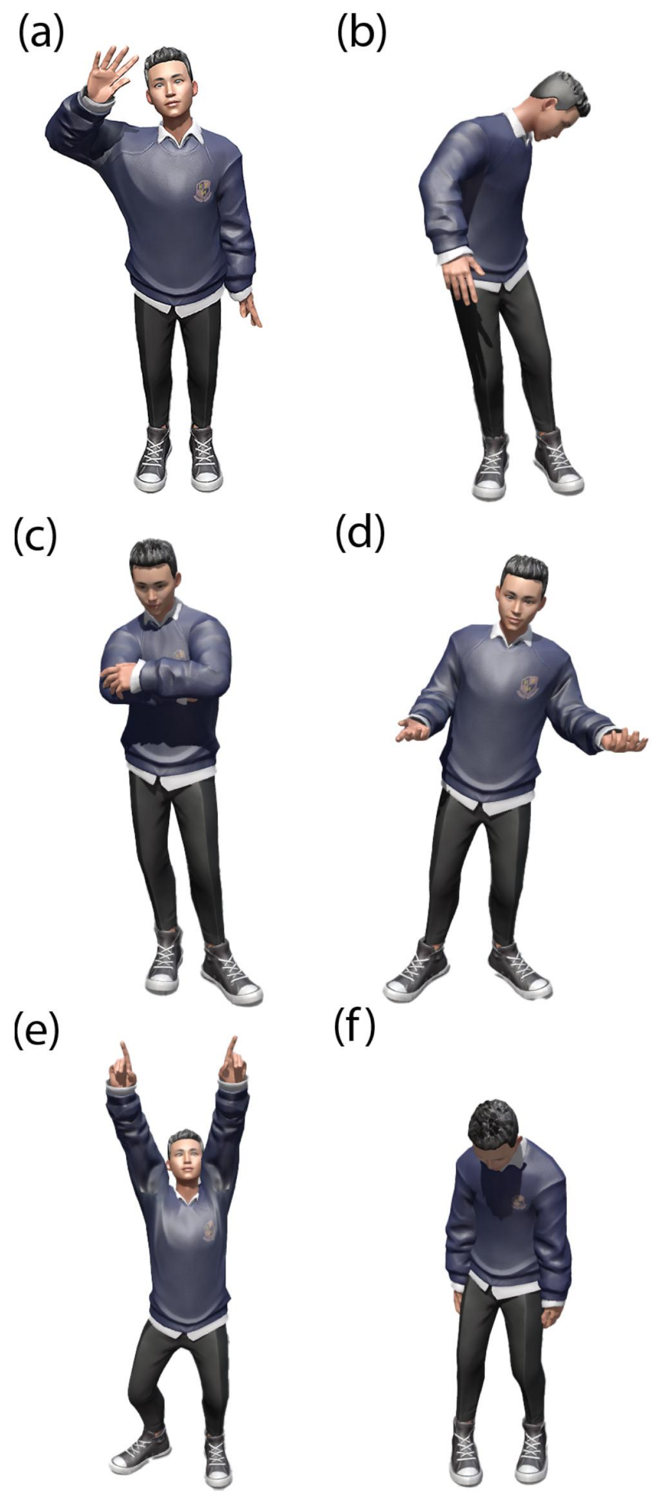


Figure 1. Examples of various postures of the S + L + P agent: a) greeting posture at game start, b) looking around to think about where to place the piece, c) folding arms and tapping the left foot tips to push users, d) neutral talking posture when the result is tie, e) victory posture to express happy emotion when the agent wins the game, and d) looking down to the ground with feet tips waving back and forth to express sad emotions when the agent loses the game.

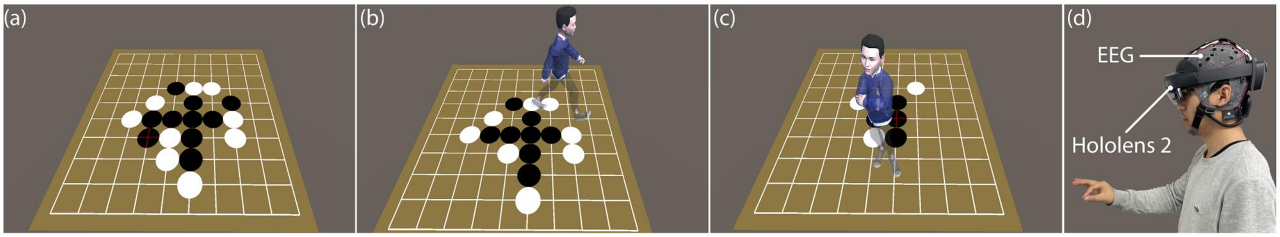


Figure 2. Three virtual agent conditions and hardware setup: a) Speech only agent, b) embodied agent with both speech and locomotion, c) embodied agent with speech, locomotion, and postures, and d) a participant wearing EEG headset and Hololens2 playing the Gobang game against the virtual agent in Mixed Reality.

3.2. MR Gobang game

The Gobang game is a turn-based game where two players play against each other, one playing white, the other black. The two players take turns to place their pieces and try to make a line of five consecutive pieces horizontally, vertically, or diagonally to win (see Li et al., 2022 for detailed rules). All possible results of this game for a player can be a win, loss, or tie. Among various board games, the Gobang game presents a unique balance between simplicity and strategic depth. This balance makes Gobang particularly suitable for examining nuanced interactions between users and IVAs without the complexities introduced by more elaborate game mechanics.

We developed an MR Gobang game based on the AlphaZero-Unity[9] and MixedRealityToolkit-Unity[10] open source projects in GitHub. The AlphaZero-Unity project trained Artificial Intelligence (AI) models based on reinforcement learning, and it also combined the Monte Carlo Tree Search (MCTS) (Wang et al., 2021) to generate strategies for the computer to play against a human player. The MixedRealityToolkit-Unity project provided support for MR interactions such as using bare hands and eye gaze tracking, and so on. In our MR Gobang game, a user wearing the Microsoft HoloLens2 played against the AI on a virtual chessboard (see section 4.3 for detailed interaction description). We captured the users' eye gaze behavior and EEG signals while they were playing the game.

3.3. Implementation

We developed the IVAs and MR Gobang game in Unity 2021.3.18f on a desktop powered by an Intel(R) Core i7-12700F CPU and NVIDIA GeForce RTX 3070 GPU. We used the Microsoft HoloLens2 to display and interact with the MR environment. To make use of the computation resources on the desktop, we utilized the Mixed Reality Toolkit 2 (MRTK2) Holographic remoting[11] mode to stream virtual content from the desktop to the Hololens2 in real-time, using a USB cable connection. We also used the Hololens2 headset to capture the user's eye gaze and gesture input as well.

For the EEG signal capture, we moved the 8 dry EEG electrodes from Unicorn wireless EEG-headset[12] to a 32 channel naked g.tech cap[13] with all original electrodes detached. The electrode layout of the 8-channel Unicorn EEG headset consisted of Fz, C3, Cz, C4, Pz, PO7, Oz and PO8. After being moved onto the g.tec headset, the relevant electrodes were AF3, P7, P3, O1, O2, P4, Pz, and Cz. We made this change because the 8 electrode positions on the Unicorn cap were not at the desired location while the g.tech cap had those locations available (see 4.5 for selecting EEG electrode positions). All of the electrodes worked at a sampling rate of 250 Hz. we used the Unicorn Suite Hybrid Black software[14] to check the connection state of each electrode and control the EEG data streaming based on the Lab Streaming Layer (LSL) library[15]. We also used the LSL in Unity to capture and store the EEG data. We added event markers to the EEG signals at specific time points, such as when the game started and stopped when the user or the virtual agent placed a piece and when the virtual agent started speaking or moving.

4. User study

In this section, we present the experimental design, set-up, task, procedure, and measurements for our user study.

4.1. Experimental design

To explore the effects of the IVAs' locomotion and postures on social presence and engagement, a within-subjects study was designed where participants played the MR Gobang game with three different agent conditions:

S: As a baseline, the virtual agent had no visual representation and used speech only (see Figure 2a)). At the beginning of the game, the virtual agent would greet the participant and invite him/her to play the Gobang game. During the game, when the participant did not play a piece for over 6 seconds, the virtual agent would encourage the participant by saying, "You are playing so slow. Please move faster" or other similar sentences. When it was the virtual agent's turn to play the chess, a black piece would be generated within one second on the board. At the end of the game, the virtual agent would speak to express emotions based on the game's result. For example, if the virtual agent won, it would be happy and say "I'm on top of the world! victory is mine!". If the virtual agent lost the game, it would express a sad emotion by saying "I'm really disappointed that I lost!". When the game was drawn, the virtual agent would be neutral about the result by saying "Well, we both gave it our best shot!". To increase the diversity of the speech content, we prepared multiple synonymous sentences for each emotional state.

S + L: Based on the S configuration, this agent also had a virtual human body standing at the center of the Gobang game board when the game started. During the game, when it came to the virtual agent's turn to play, it would walk towards the target position before a new black piece was generated, as is shown in Figure 2(b). Once the virtual agent stopped walking, it would turn to face the participants and talk to them. The virtual agent's body was semi-transparent to allow the player to place pieces behind it if needed.

S + L+P: Based on the S + L condition, the agent also had postures to express thinking behavior and emotions based on the game state. For example, during the game, the virtual agent could look around the board before starting to walk as if thinking about where to place the next piece. While the virtual agent was encouraging the participant to play, it would have two arms folded and the left foot tip-tapping on the ground to show its impatience (see Figure 2(c)). At the end of the game, if the virtual agent won, it would show a victory posture while speaking to express happiness. When the virtual agent lost the game, it would show sad emotion by looking down and kicking the ground slowly but repeatedly with the right leg. However, when the result was drawn, the virtual agent would just show neutral emotion with two hands waving naturally in front of the body while talking.

We collected data from 16 participants (10 male and 6 female) with ages ranging from 20 to 33 years old ($M = 27.56$, $SD = 3.56$) from a university campus. Participants were recruited based on their familiarity with English and their willingness to wear both an EEG cap and an MR headset. Eight of them reported they had no experience in using their hands to interact with virtual content in AR/VR/MR, seven of them reported that they had little experience (i.e., semesterly/annually), and one participant reported that he often uses this interaction mode (i.e., daily/weekly/monthly). Ten rated themselves as a novice for the Gobang game, followed by four people at the intermediate level and two as advanced. None rated themselves as experts or players who had never played the game. Eight of them reported that they had little experience interacting with IVAs like Amazon Alexa, Microsoft Cortana, Apple Siri, etc. Five people reported they often used such agents, while three said that they had never used any agents like this. We used a balanced Latin Square to counterbalance the order of conditions and reduce the learning effect.

In the user study, we mainly investigated the following research question: *How do the virtual agent locomotion and posture influence the user's social perception and engagement?*

To address this research question, we formulated two main research hypotheses, each with two sub-hypotheses, based on our experimental design outlined above. Our research hypotheses were as follows:

H1a: Compared to a speech-only virtual opponent, incorporating locomotion in the virtual agent can enhance human-agent social interaction as indicated by the subjective and behavioral measurements presented in Section 4.5.1 and 4.5.3, such as social presence, engagement, gaze duration proportion, and task performance.

H1b: Combining locomotion and postures for the virtual agent can further improve the social perception as demonstrated by the measurements presented in Section 4.5.1 and 4.5.3.

H2a: Compared to a speech-only virtual opponent, incorporating locomotion in the virtual agent can increase the EEG-based engagement index (see Section 4.5.2).

H2b: Combining locomotion and postures for the virtual agent can further increase the EEG-based engagement index (see Section 4.5.2).

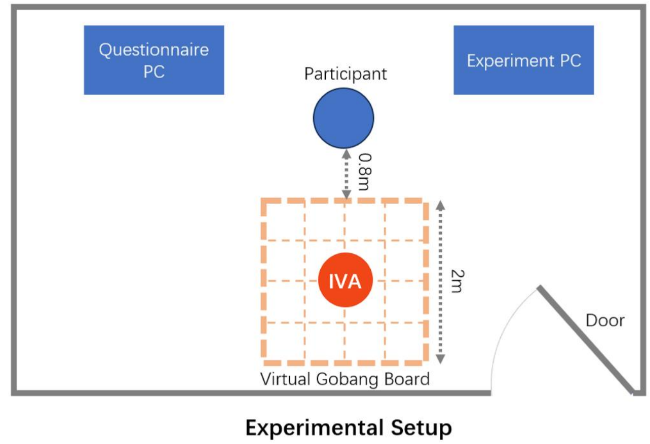


Figure 3. Experimental setup. A participant with EEG and Hololens 2 headset stands facing the wall. Through the Hololens headset, the participant could see a 2 m*2m virtual square board on the floor with a virtual human standing at the center facing the participant. The nearest edge of the chessboard to the participant was around 0.8 m away.

We formulated H1, including H1a and H1b, based on insights from the literature review (see Section 2.1). For H2, including H2a and H2b, we hypothesized that if H1 was confirmed in our experimental design, corresponding significant effects would also be observed in the EEG signals.

4.2. Experimental setup

The experiment was conducted in an isolated experiment room. As shown in Figure 3, the participant wore an EEG cap and Hololens 2 headset (see Figure 2(d)) and faced the wall in front of the experiment PC, which was used for running the MR Gobang game application. Through the Hololens headset, the participant could see a 2 m*2m virtual square board lying on the floor with a virtual human standing at the center and facing the participant. The nearest edge of the chessboard to the participant was around 0.8 m away. A second Questionnaire PC was used in the room corner to fill in questionnaires after each session.

4.3. Experimental task

In the experiment, the participants played the MR Gobang game against different virtual agents under different study conditions. As shown in Figure 4, after the program started, the virtual agent began interacting with the participants by

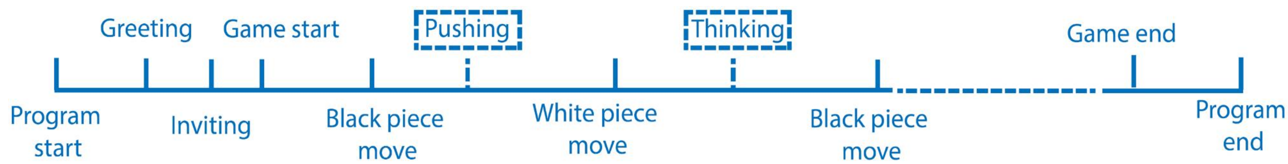


Figure 4. Events in the game process. Each bar represents one event, and these events happen one after another from left to right. The distance between each other doesn't mean the time length. The "pushing" and "thinking" are highlighted with dashed lines, representing these two types of events are not always triggered. Except for "program start" and "program end", other events were annotated as markers in EEG recording.

greeting them verbally. Then the virtual agent would invite the participants to play the Gobang game and ask whether they were ready. Once the participants verbally replied to the virtual agent, the game started, triggering the EEG data recording and eye gaze behavior logging functions.

During the game, the virtual agents always played the black piece and moved first while the participants played the white piece. The participant controlled a virtual beam shot from the center of their palm to select the target position on the board where they wanted to place their piece. When the beam hits the board, there is a circle cursor to indicate the position. Once the cursor was moved close to a cross-dot on the chessboard, the cross-dot would be highlighted with a virtual rectangle to indicate the target dot's selection state. Once the participant decides to put a white piece at the currently selected dot, he/she only needs to tap with his/her thumb and index fingers. The user's goal was to win against the virtual agent.

Once the game finished, the virtual agent responded to the result based on its designed features (see section 4.1). After the virtual agent's response, the game ended.

4.4. Experimental procedure

Once the participants entered the experiment room, they were guided to sit in front of the questionnaire PC to answer questions and were provided with the consent form. Once they agreed to participate in the experiment, they were asked to switch off all their electronic devices, such as phones and smartwatches, to reduce interference of the EEG signal during the experiment. They were asked to fill in a demographic questionnaire. Then, they were placed in another chair facing a wall. This chair was placed close to the experiment PC to run the MR Gobang game. Then the participants donned a Hololens2 headset and went through the eye calibration procedure to calibrate their eye gaze with the headset. After the eye gaze calibration, the participants were given three training opportunities to play the MR Gobang game. This was done without showing any agent cues to get familiar with basic game rules and using the dominant bare hand to place virtual pieces. They were asked to stand up before playing the game.

After the training, the experimenter helped the participants put on the EEG headset, followed by the Hololens2 headset. We checked the EEG electrode connection state and the EEG signals (see section 3.3). When the EEG signals stabilized, we started the experimental task where the participant was asked to play against a virtual agent in the MR Gobang game.

Before the experimental task, the participants were asked to stand still with their eyes closed and heads facing forward. As they heard the virtual agent's greeting voice, the task started (see section 4.3 for a detailed task description). After each task session, the participants were asked to fill out four questionnaires as presented in section 4.5.1. Participants were given a 1-minute rest after completing the questionnaires before starting the next condition.

Each participant took three trials in total, and the whole experiment lasted around 80 minutes. After the experiment, a semi-structured interview was conducted to understand the participants' experience and perceptions of the virtual agent through the user study.

4.5. Measurements

4.5.1. Subjective measurements

Social engagement. Inspired by research in Human-Robot Interaction (HRI) that explores engagement between users and social robots (Jung et al., 2023; Riedmann et al., 2024), we were also interested in how our different virtual agent designs influence users engagement with the virtual agents. As presented in Sidner et al. (2004), engagement is "The process by which two (or more) participants establish, maintain and end their perceived connection." Later, Corrigan et al. (2013) further identified this type of engagement as social engagement. In our work, we used the questionnaire from Leite et al. (2014), which was developed based on the social engagement concept to measure the user's social engagement with the virtual agent.

Social presence. To understand the impact of virtual agent locomotion and postures on users' feelings of social presence with the agent, we measured the social presence using questionnaire presented in Harms and Biocca (2004), which consists of six subscales: *co-presence*, *attentional allocation*, *perceived message understanding*, *perceived affective understanding*, *perceived affective interdependence*, and *perceived behavioral interdependence*.

Emotional arousal and valence. We further examined the potential influence of the embodied virtual human's locomotion and postures on users' emotional responses. So we used the Self-Assessment Manikin (SAM) Valence Arousal questionnaire (Lang, 1995) to measure users' emotional arousal and valence.

NASA-TLX weighted workload. We also used the NASA-TLX questionnaire (Hart & Staveland, 1988) to assess users'

Table 1. Interview questions and motivations.

No.	Interview question	Motivation
Q1	Which agent did you like the most, and why?	User preference
Q2	Which agent did you hate the most, and why?	User preference
Q3	In which condition did you perceive the strongest emotional cues?	Emotion perception
Q4	Do you prefer competing against the agent or having the virtual agent assist you in the game?	Agent role preference

overall workload while they played the Gobang game against different virtual agents.

Interview questions. We conducted a short semi-structured interview after all the experiment condition trials to understand the participants' experience and perceptions of the virtual agent through the user study. The interview questions and motivations are shown in Table 1. The Q1 and Q2 were designed to understand the user's preference for different agent conditions. The Q3 was designed to understand the user's perception of the virtual agent's emotional cues. Although Q4 was not directly related to the primary research questions, it was designed to gauge participants' preferences regarding the agent's role in the game. In the current experiment, the agent acted as an opponent. To clarify the assistant role for participants, we explained that the agent would have the same appearance and could also walk on the chessboard to suggest positions when asked. We expected the answers to Q4 could help us figure out the potential relationship between users' skill level and their preference for the agent's role in the game, which could be useful for future research.

4.5.2. Physiological measurements

EEG-based engagement index. As illustrated in Figure 5, we collected the raw EEG data from the AF3, P7, P3, O1, O2, P4, Pz, and Cz electrodes. The P3, P4, Pz, and Cz were selected according to the indices of operator engagement introduced in Pope et al. (1995), while the AF3 and P7 electrodes were chosen because these two electrodes were shown effective in capturing user engagement in a video game study (Ruqeyya et al., 2022). Furthermore, the O1 and O2 electrodes were also selected because the occipital lobe was related to vision processing (Malach et al., 1995). As shown in Section 2.3, we used the EEG engagement index formula recommended by Pope et al. (1995) to check the participants' cognitive engagement during the game. The formula is as follows:

$$\text{EEG-based engagement index} = \frac{\beta}{\alpha + \theta} \tag{1}$$

The β, α, and θ bands were originally combined in Pope et al. (1995) by summing the relative frequency band power calculated from the Cz, Pz, P3, and P4 electrodes. In addition to this combined approach, other studies have calculated the EEG-based engagement index using single-electrode data from various brain regions, such as the frontal, occipital, parietal, and temporal lobes (Alimardani et al., 2021). In our study, we examined the EEG-based engagement index using both combined electrode data and single-

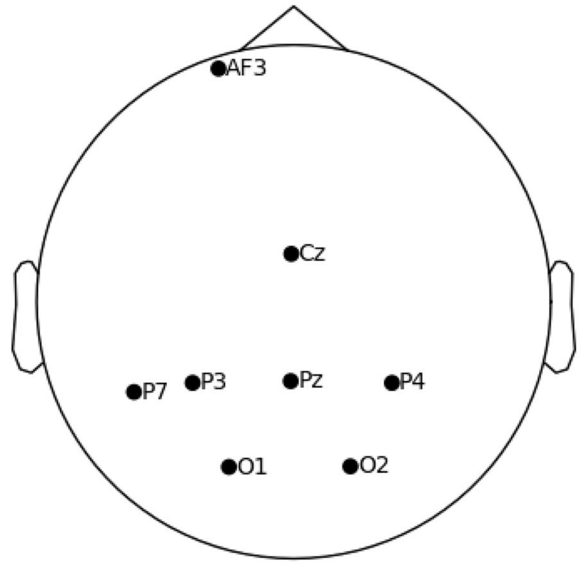


Figure 5. EEG sensor positions. We calculated the EEG-based engagement index using data from these sensor positions (Alimardani et al., 2021; Pope et al., 1995).

electrode data to explore the differences in engagement between different agent conditions.

Artifacts in the EEG data can be caused by wires, eye blinking, muscle movement, and electrode movement (Chang et al., 2022; Ruqeyya et al., 2022). To reduce these, we first applied a bandpass filter (1∼40Hz), followed by manually checking each participant's EEG signals per condition and removing bad data segments using the MNE-Python (Gramfort et al., 2013). In total, 6.84% of the data was identified as bad and subsequently removed. Then, we chose the 1-second data before the "Greeting" event marker (see Figure 4) as a baseline and applied baseline correction[16] to normalized the raw EEG. After the baseline correction, we chunked the data between the "Greeting" marker and the "Game end" marker into 1-second length epochs, with which we further calculated the Power Spectral Density (PSD) for each frequency band in the equation 1 using the Welch method (Welch, 1967). We finally averaged the PSD values for each frequency band across all epochs to calculate the EEG-based engagement index for each participant and each condition. Due to technical reasons, three participant's EEG data (2 male, and 1 female) was not recorded correctly and excluded from the EEG data processing.

4.5.3. Behavioral measurements

Gaze duration (GD) proportion. To further understand the user's visual attention allocation in the S + L and S + L + P conditions, we measured the user's eye gaze behavior by logging the start time (t_{start}) and stop time (t_{stop}) when users

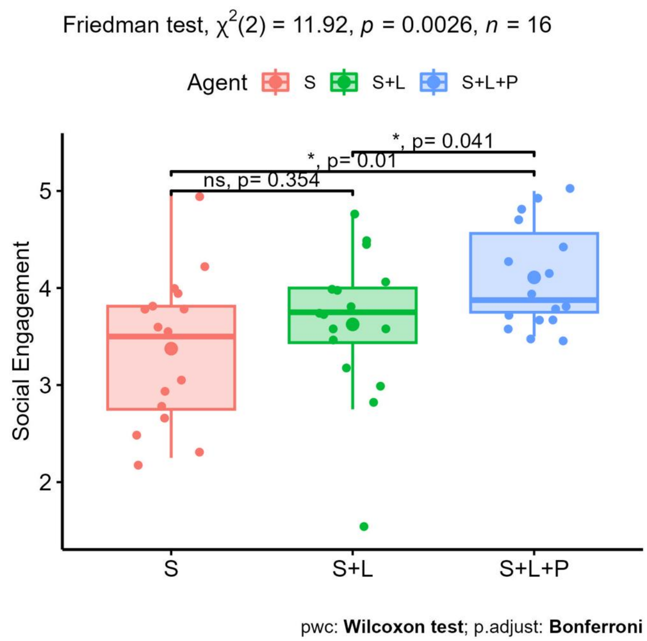


Figure 6. Results of the engagement questionnaire. (5-point Likert Scale from 1 to 5; the larger solid dot in each condition represents mean value; Statistical significance: * ($p < 0.05$)).

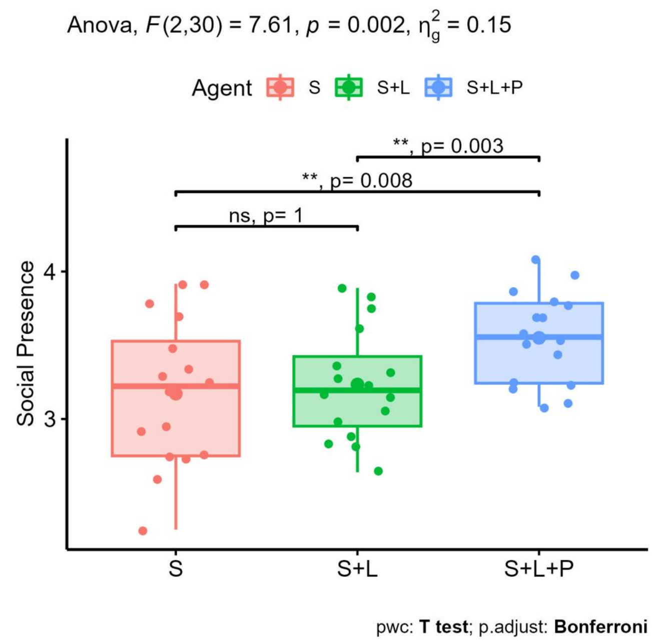


Figure 7. Results of the social presence. (5-point Likert Scale from 1 to 5; the larger solid dot in each condition represents mean value; Statistical significance: * ($p < 0.05$), ** ($p < 0.001$)).

looked at either the board ($GD_{chessboard}$) or agent body separately. Based on the GD_{agent} and $GD_{chessboard}$, we further calculated the agent GD proportion (see formula 2), which means the proportion of time duration participants gazed at the virtual human body. For the agent GD proportion, we didn't consider the S condition because the virtual agent didn't have a virtual body in this condition.

$$\text{Gazing Duration (GD)} = t_{stop} - t_{start} \tag{2}$$

$$\text{Agent GD proportion} = \frac{\sum_{i=1}^{n} GD_{agent}}{\sum_{i=1}^{n} GD_{agent} + \sum_{i=1}^{n} GD_{chessboard}} \times 100\% \tag{3}$$

where i means the participants index

Player average speed. We measured the player's average speed by logging the start time when the game transitioned to the participant's turn and the stop time when the participant completed their move. Similar to the intention of capturing users' eye gaze behaviors, we analyzed the player's average speed to examine whether the virtual agent's body attracted more attention, potentially leading to a longer time for each move.

Rounds of play. We also logged the number of rounds of play to understand how the virtual agent's behavior influenced the overall game experience. As described in Section 4.3, the agent always moved first, so the number of rounds of play was the same as the number of moves the participant made.

Task performance. Besides capturing user attention allocation as measured by the eye gaze behavior, player average speed, and task completion time, we also logged the task performance to understand how well participants performed in the game under different agent conditions. The task performance result could be a user win, user loss, or tie.

Task completion time. We logged the task completion time to understand how the virtual agent behavior influenced the overall task completion time. The task completion time was calculated as the time from the game start to the game end.

5. Results

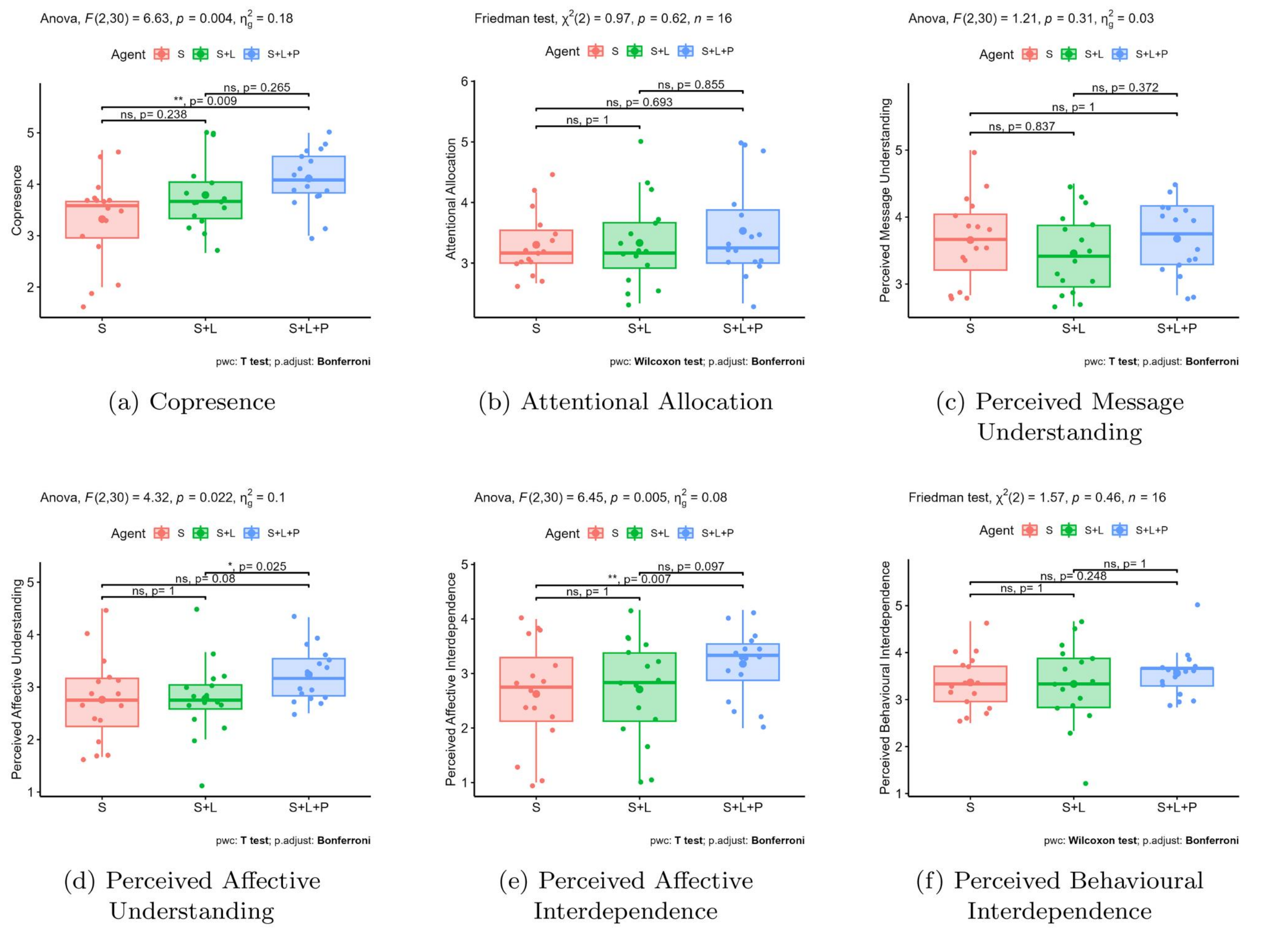
5.1. Subjective measurements results

5.1.1. Social engagement

The Shapiro-Wilk test showed the engagement questionnaire data violated the normality assumption in the S + L + P ($p = 0.031$) condition. As shown in Figure 6, the Friedman test indicated there was a significant difference among the three agent conditions ($\chi^2(2) = 11.92, p = 0.0026$) with moderate effect size ($W = 0.372$). The pairwise Wilcoxon signed-rank tests with Bonferroni correction revealed that the engagement in the S + L + P was significantly higher than that in the S condition ($Z = -2.576$, $p = 0.01$) and the S + L condition ($Z = -2.044$, $p = 0.041$). No other significant difference was found in other pairwise comparisons.

5.1.2. Social presence

The Shapiro-Wilk test showed that the social presence data in S ($p = 0.648$), $S + L(p = 0.541)$ and S + L + P ($p = 0.580$) were all normally distributed. Mauchly's test indicated the sphericity assumption was also met ($W = 0.739$, $p = 0.12$). As shown in Figure 7, we saw overall a significant difference across the three experimental conditions using the repeated ANOVA ($F(2, 30) = 7.607, p = 0.002, \eta_g^2 = 0.15$). The Post-Hoc pairwise t-tests with Bonferroni correction showed that



(a) Copresence (b) Attentional Allocation (c) Perceived Message Understanding

(d) Perceived Affective Understanding (e) Perceived Affective Interdependence (f) Perceived Behavioural Interdependence

Figure 8. Boxplots of subscales from social presence questionnaire (5-point Likert Scale from 1 to 5; the larger solid dot in each condition represents mean value; Statistical significance: * ($p < 0.05$), ** ($p < 0.001$).

the social presence in the S + L + P condition was significantly higher than that in the S condition ($t(15) = -2.652, p = 0.008$) and S + L condition ($t(15) = -2.958$, $p = 0.003$). No significant difference was found between the S and S + L conditions.

Copresence. The Shapiro-Wilk test showed the copresence data in S ($p = 0.09$), $S + L (p = 0.098)$ and S + L + P ($p = 0.578$) were all normally distributed. Mauchly's test indicated the sphericity assumption was met ($W = 0.86$, $p = 0.348$). As shown in Figure 8(a), the repeated ANOVA revealed a significant difference among the three agent conditions ($F(2, 30) = 6.633$, $p = 0.004$, $\eta_g^2 = 0.176$). The Post-Hoc pairwise t-tests with Bonferroni correction revealed that the copresence in the S + L + P condition was significantly higher than that in the S condition ($t(15) = -2.612$, $p = 0.009$). No other significant difference was found in other pairwise comparisons.

Attentional allocation. The Shapiro-Wilk test indicated the normality assumption was violated in the attentional allocation data in S + L + P ($p = 0.03$). No significant difference was found among the three conditions using the Friedman test (see Figure 8(b)).

Perceived message understanding. The Shapiro-Wilk test revealed the perceived message understanding data was normally distributed in S ($p = 0.329$), $S + L (p = 0.306)$ and S + L + P ($p = 0.126$). Mauchly's test indicated the sphericity assumption was met ($W = 0.90$, $p = 0.478$). We couldn't find a significant difference in the perceived message understanding subscale of social presence among the three agent conditions using the repeated ANOVA test (see Figure 8(c)).

Perceived affective understanding. The Shapiro-Wilk test showed the perceived affective understanding data was normally distributed in S ($p = 0.477$), S + L ($p = 0.373$), and S + L + P ($p = 0.427$). Mauchly's test indicated the sphericity assumption was also met ($W = 0.889$, $p = 0.438$). As shown in the Figure 8(d), the repeated ANOVA revealed a significant difference among the three agent conditions ($F(2, 30) = 4.319$, $p = 0.022$, $\eta_g^2 = 0.095$). The Post-Hoc pairwise t-tests with Bonferroni correction showed that the perceived affective understanding in the S + L + P condition was significantly higher than that in the S + L condition ($t(15) = -2.241$, $p = 0.025$). No other significant difference was found in other pairwise comparisons.

Perceived affective interdependence. The Shapiro-Wilk test indicated the perceived affective interdependence data was

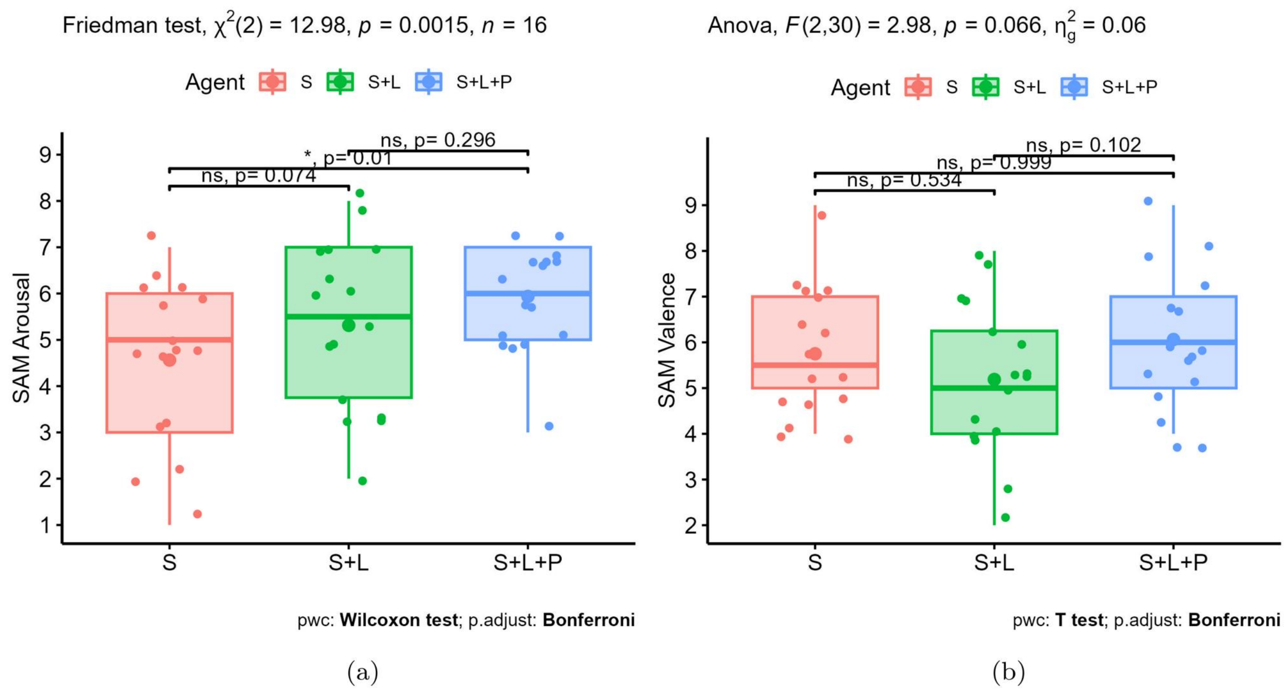


Figure 9. Boxplots of self Assessment Manikin (SAM) arousal Valence questionnaire (9-point Likert Scale from 1 to 9; the larger solid dot in each condition represents mean value; Statistical significance: * ($p < 0.05$).

normally distributed in S ($p = 0.357$), S + L ($p = 0.415$) and S + L + P ($p = 0.285$). Mauchly's test indicated the sphericity assumption was met ($W = 0.808$, $p = 0.224$). As illustrated in the Figure 8(e), the repeated ANOVA revealed a significant difference among the three agent conditions ($F(2, 30) = 6.45$, $p = 0.005$, $\eta_g^2 = 0.078$). The Post-Hoc pairwise t-tests with Bonferroni correction showed that the perceived affective interdependence in the S + L + P condition was significantly higher than that in the S + L condition ($t(15) = -2.697$, $p = 0.007$). No other significant difference was found in other pairwise comparisons.

Perceived behavioural interdependence. The Shapiro-Wilk test showed the perceived behavioral interdependence data violated the normality assumption in the condition S + L + P ($p = 0.026$). No significant difference was found among the three conditions using the Friedman test (see Figure 8(f)).

5.1.3. SAM valence-arousal scale

For the SAM Arousal scale, the Shapiro-Wilk test indicated the SAM Arousal data in S ($p = 0.032$) and S + L + P ($p = 0.004$) violated the normality assumption. As shown in Figure 9(a), the Friedman test showed there was a significant difference among the three agent conditions ($\chi^2(2) = 13.0, p = 0.002$), with a moderate effect size ($W = 0.406$) by the following Kendall's W test. The post hoc Wilcoxon signed-rank test with Bonferroni correction revealed the arousal in the S + L + P condition was significantly higher than that in the S condition ($Z = -2.576$, $p = 0.01$). No significant difference was found in other pairwise comparisons.

For the SAM Valence scale, the Shapiro-Wilk test showed the normality assumption was met. Mauchly's test ($W = 0.926$, $p = 0.584$) did not indicate any violation of sphericity. No significant difference was found using the repeated ANOVA among the three conditions ($F(2, 30) = 2.976$, $p = 0.066$, $\eta_g^2 = 0.055$).

5.1.4. NASA-TLX weighted workload

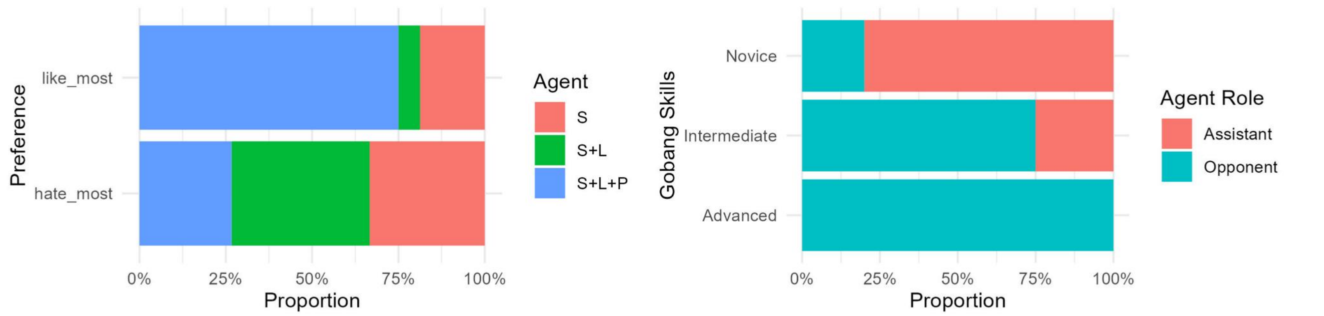
The Shapiro-Wilk normality test showed that all NASA-TLX weighted workload data of S (p =.801), S + L (p =.725), S + L + P (p =.064) met the normality assumption. Mauchly's test showed the NASA-TLX weighted workload data also met the sphericity assumption ($W = 0.856$, $p = 0.363$). No significant difference was found among the three agent conditions using a repeated ANOVA ($F(2, 30) = 0.370$, $p = 0.694$).

5.1.5. Interview on user preference

After interacting with all three agent conditions, participants ranked the conditions they most and least liked. A Chi-Square Goodness of Fit test yielded a significant difference against the most liked agent type (χ^2 (2) = 19.313, $p < 0.0001$). However, no significant difference was found on the most hated agent type (see Figure 10).

5.1.6. Interview on emotional cues perception

All the participants reported that they perceived the strongest emotional cues in the S + L + P condition. Detailed comments are quoted in the discussion section to support the discussion.



(a) (b)

Figure 10. Agent behavior and role preference.

5.1.7. Interview on agent role preference

A Chi-squared test of independence revealed a significant association between Gobang skill level and agent role preference (χ^2 (2) = 6.451, $p = 0.04$). As shown in Figure 10, participants with higher Gobang skill levels preferred the opponent role, while participants with lower skill levels preferred the assistant role.

5.2. Physiological measurement results

5.2.1. EEG-based engagement index

The Shapiro-Wilk normality test on the EEG-based engagement index indicated the normality assumption was violated in all single electrodes. As shown in the Figure 11, the Friedman test showed there was a significant difference among three agent conditions $(\chi^2(2) = 6.62, p = 0.037)$ at AF3. The follow-up Kendall's W test revealed a small effect size ($W = 0.254$). The post-hoc pairwise Wilcoxon signed-rank tests with Bonferroni correction revealed that the EEG-based engagement index in the S + L + P group was significantly lower than those in the S + L ($Z = -2.366$, $p = 0.018$). No other significant difference was found regarding the EEG-based engagement index.

Since the EEG-based engagement index and social engagement questionnaire (see Section 5.1) were both used to measure the user's engagement, we further analyzed the correlation between the social engagement and EEG-based engagement index. The Pearson correlation test showed no significant correlation between the social engagement and EEG-based engagement index in the S ($r = -0.140, p = 0.648$), $S + L(r = -0.216, p = 0.479)$ *and* $S + L + P(r = -0.172, p = 0.574)$ conditions.

5.3. Behavioral measurements results

5.3.1. GD proportion

The Shapiro-Wilk test indicated that eye gaze duration proportion in both the S + L ($p = 0.411$) and S + L + P ($p = 0.124$) conditions were normally distributed. No significant difference was found between these two groups using a pairwise t-test with Bonferroni correction ($t(15) = -0.951$, $p = 0.357$). Descriptive statistics showed the S + L + P ($M = 51.204\%$, $SD = 13.163\%$) had higher eye duration proportion than the S + L ($M = 47.055\%$, $SD = 12.162\%$).

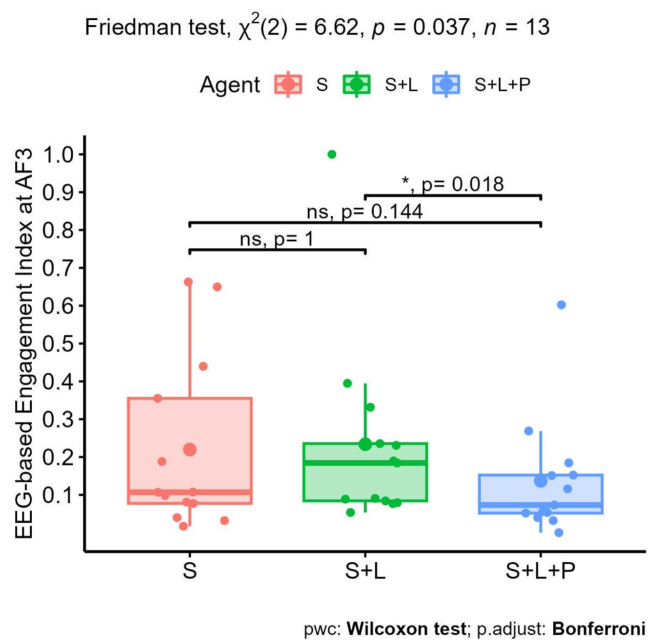


Figure 11. Results of the EEG-based engagement index at AF3(three participants' data was not correctly recorded and excluded from EEG data processing; the larger solid dot in each condition represents mean value; Statistical significance: **** ($p < 0.0001$).

5.3.2. Player average speed

The Shapiro-Wilk normality test showed that the player average speed of S ($p = .120$), S + L ($p = .186$), and S + L + P ($p = .162$) all met the normality assumption. The Mauchly test showed that the sphericity assumption was met ($W = 0.832$, $p = 0.303$). No significant difference was found across the three conditions by using a repeated measure ANOVA with sphericity assumed ($F(2, 28) = 0.218$, $p = 0.805$). Descriptive statistics indicated that participants in the S + L + P condition ($M = 6.21$, $SD = 2.21$) and the S + L condition ($M = 6.45$, $SD = 2.30$) played slightly slower, taking longer to complete each move, compared to the S condition ($M = 6.16$, $SD = 2.32$).

5.3.3. Rounds of play

The Shapiro-Wilk normality test showed that the rounds of play in S ($p = 0.021$), and S + L + P ($p < 0.001$) violated the normality assumption. No significant difference was found across the three conditions by using a Friedman test

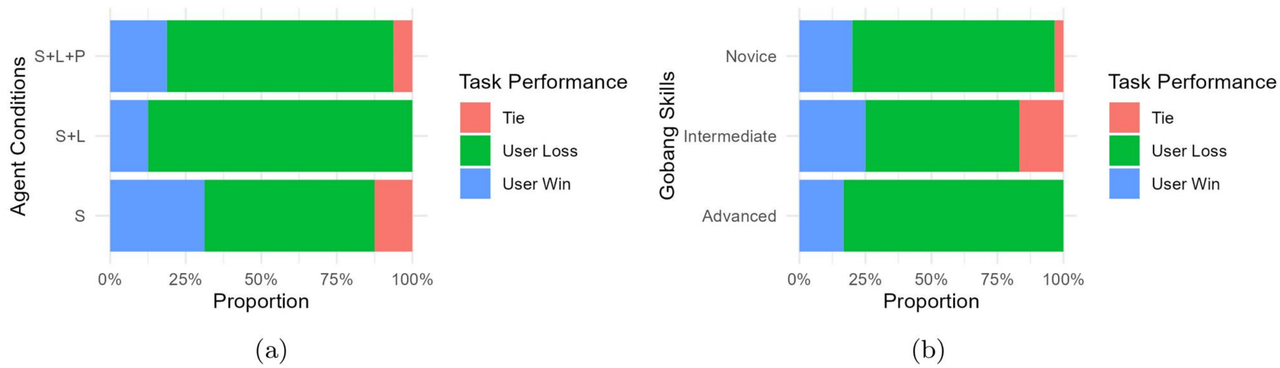


Figure 12. Analysis of task performance in relation to agent conditions and participant Gobang skills.

($\chi^2(2) = 3.12, p = 0.21$). Descriptive statistics showed that participants in the S + L + P condition ($M = 16$, $SD = 11.6$) and the S + L condition ($M = 15$, $SD = 7.28$) played fewer rounds compared to the S condition($M = 24$, $SD = 13.9$).

5.3.4. Task performance

Task performance was measured based on game outcomes, categorized as user win, user loss, or tie, making it a categorical variable. A Chi-square test of independence was conducted to examine the relationship between task performance and agent conditions. As shown in Figure 12 (a), the results showed no significant association between task performance and the agent conditions (χ^2 (4) = 4.486, $p = 0.344$). Interestingly, when applying Chi-square tests to the three different game outcomes within each agent condition to test the Goodness of fit, significant differences were found between the game outcomes in the S + L condition (χ^2 (1) = 9, $p = 0.003$) and the S + L + P condition (χ^2 (2) = 12.875, $p = 0.002$), while no significant difference was observed in the S condition (χ^2 (2) = 4.625, $p = 0.099$).

Moreover, considering that participants' experience in the Gobang game might influence the game outcomes, we further examined the relationship between task performance and user Gobang skills, as reported in the demographic questionnaire. A Chi-square test of independence was conducted to assess this relationship. As shown in Figure 12(b), the results indicated no significant association between task performance and user Gobang skills (χ^2 (4) = 3.524, $p = 0.474$).

5.3.5. Task completion time

The Shapiro-Wilk test showed the normality assumption on the task completion time was violated in the S + L + P condition ($p < .001$). No significant difference was found in the task completion time over the three conditions using a Friedman test ($\chi^2(2) = 0.875, p = 0.646$). Descriptive statistics showed the S + L + P ($M = 178.88$, $SD = 118.72$) and S + L ($M = 148.31$, $SD = 91.72$) conditions had shorter task completion time than the S ($M = 186.25$, $SD = 118.72$) condition.

6. Discussion

In general, we found significant differences between the three agent conditions in subjective and physiological measurements, while no significant differences were observed in the behavioral measurements. Notably, all significant differences occurred between the S + L + P condition and the other two conditions, with no significant differences between the S and S + L conditions. Thus, our results did not support H1a or H2a. However, the subjective measurements supported H1b, while the EEG-based engagement index contradicted H2b. The discrepancy between the subjective measurements and the EEG-based engagement index may be due to the different aspects of engagement each method captures, and no correlation was found between the two measures in our results (see Section 5.2). The following sections will discuss these findings in more detail.

6.1. Combining locomotion and postures enhances social perception of the Mixed Reality agent

Our subjective questionnaire results of social presence, social engagement, SAM arousal, and interview on user preference supported H1b but rejected H1a.

Participants reported stronger overall social presence and engagement while playing against the S + L + P agent compared with the S agent. For the S + L + P agent, the virtual agent locomotion behavior in the MR environment and postures like folding arms to push participants, looking around the board as if it was thinking about where to place the piece, and expressing emotional postures at the game end, exerting social cues into the game and thus improved the social presence and engagement. These results are in line with previous work (Kim et al., 2018b), where they also argued that the virtual agent's appropriate social behavior played a role in improving social presence and engagement.

In addition to the overall social presence and engagement, we further examined the subscales of social presence and found that the S + L + P agent exhibited significantly higher copresence and perceived affective interdependence compared to the S condition, as well as significantly higher perceived affective understanding compared to the S + L condition. These results could be validated by the participants' comments towards interview Q1. For example, P19 commented *"I like the agent with posture because he expressed emotional postures making me feel as if interacting with a real person"* and P08 commented *"I think the agent with posture looks more alive and the virtual body could sometimes remind me about where the black piece was placed."* Moreover, P07 reported that *"… overall I could feel the virtual agents were stronger than me* *in playing this game. But the agent with postures like looking around before moving his pieces made me more positive and confident because I thought it looked like he also needed to think hard to win me…"* These findings suggest that the S + L + P agent was perceived as more alive, interactive, and emotionally expressive, enhancing participants' social perception of the virtual agent.

Regarding perceived emotional cues, our SAM arousal results indicated that the S + L + P agent elicited significantly higher emotional arousal compared to the S agent. This finding was consistent with the social presence subscale results, specifically in perceived affective interdependence and perceived affective understanding. Furthermore, in interview Q3, all participants reported perceiving the strongest emotional cues in the S + L + P condition, further supporting the effectiveness of this condition in enhancing users' social perceptions.

However, no significant difference was found in the SAM valence, though the SAM arousal was significantly higher in S + L + P compared to S conditions. This could be because, in general, participants perceived stronger emotional cues in the S + L + P agent, but some people felt positive towards the perceived emotional cues while others felt negative. For example, P04 commented while answering the interview Q1, "*… the agent with postures was more interactive and likable…*" On the contrary, P10 commented while answering the interview Q2, "*I hate the virtual agent showing postures because his behavior of folding up his arms and asking me to move faster made me think he was impolite. However, while interacting with the speech-only agent, I could not tell if he was pushing me politely or not…*" This could partially explain why the SAM valence didn't show a significant difference in S + L + P compared with the S conditions.

Lastly, our findings did not support H1a, which may be because that providing the embodied virtual agent with locomotion alone did not significantly enhance users' social perceptions of the agent in this experiment. It is possible that the agent's locomotion behavior was less effective than postures in fostering social presence and engagement. In interview Q2, P15 remarked, "*… the locomotion agent didn't make sense to me and at times its movement distracted my attention…*", while P17 noted, "*… the locomotion agent sometimes stared at me, which felt cold and rigid…*" These comments suggest that locomotion may not be as impactful as postures in promoting social presence and engagement.

Overall, our subjective results suggested that combining locomotion and postures in the virtual agent could enhance social presence and engagement, as well as perceived emotional cues, in the MR environment. The S + L + P agent was perceived as more alive, interactive, and emotionally expressive, which may have contributed to the enhanced social perception of the agent.

6.2. Increased social perception in Mixed Reality agent may lead to a lower EEG-based engagement index in the board game

Our EEG engagement index results led to the rejection of both H2a and H2b. No significant differences were observed in the EEG-based engagement index, either at individual electrodes (except AF3) or across combined electrodes, between the S + L and S conditions, thus rejecting H2a. Furthermore, the engagement index at electrode AF3 was significantly lower in the S + L + P condition compared to the S + L condition, which directly contradicts H2b. No other significant differences were found between the S + L and S + L + P conditions, further supporting the rejection of H2b.

Interestingly, while the EEG-based engagement index at AF3 was significantly lower in the S + L + P condition, this was opposite to the social engagement questionnaire results. However, no significant correlation was found between the EEG engagement index and the social engagement questionnaire scores. This discrepancy may be due to the different aspects of engagement each method captures. The engagement questionnaire mainly focused on how participants interacted with the virtual agent, as the statements are specifically related to human-agent interaction (Leite et al., 2014). In contrast, the EEG engagement index measured participants' information-gathering, visual processing, and allocation of attention (Berka et al., 2007; Pope et al., 1995). Furthermore, the electrode AF3 captured the brain activities around the anterior-frontal cortex, which was associated with attention and cognitive control Chun et al. (2011).

In our study, participants were primarily focused on the Gobang game, likely dividing their attention between interacting with the virtual agent and concentrating on game strategy. As shown by the descriptive results of gaze proportion (see Section 5.3), in the S + L + P condition, participants spent over 50% of their gaze duration on the virtual agent, while in the S + L condition, nearly 50% of their gaze duration was directed at the agent. This suggested that the virtual agent's body and behaviors like locomotion and postures drew approximately half of the participants' visual attention away from the game board. This finding aligns with participants' comments during interview Q2. For example, P10 stated, "I hate the virtual agent showing postures because his behavior of folding up his arms and asking me to move faster made me think he was impolite. However, while interacting with the speech-only agent, I could not tell if he was pushing me politely or not." And P19 said, "… the virtual agent's body made no sense to me and sometimes even occluded my target on the chessboard…" These comments underscore that the agent's physical movements and presence were perceived as distracting and, in some cases, obstructive to the gameplay experience.

Consistent with the descriptive statistics of gaze proportion, other behavioral measures, such as average player speed, rounds of play, and task completion time, showed that the embodied agent conditions, S + L and S + L + P, were associated with slower player speeds, fewer rounds of play, and shorter task completion times compared to the S condition. Additionally, our task performance results indicated that participants in the S + L and S + L + P conditions experienced significantly more losses than wins or ties, while no significant performance distribution was observed in the

S condition. Importantly, participants' self-reported Gobang skill levels did not significantly influence task performance and the different agent behaviors did not significantly affect workload, as evidenced by the NASA TLX results. These findings suggested that the virtual agent was generally more skilled than most participants, and the embodied agent's locomotion and behaviors might have distracted participants from the game's logic and strategy, potentially resulting in the higher loss rates in the S + L and S + L + P conditions.

In particular, the social behaviors of the S + L + P agent, such as looking around the board before making a move, folding its arms to prompt participants, and expressing emotional postures, might have drawn participants' attention away from the game, resulting in slower gameplay and fewer rounds. This divided attention could also explain why the EEG engagement index at AF3 was significantly lower in the S + L + P condition compared to the S + L condition, as participants were splitting their focus between interacting with the virtual agent and concentrating on the game.

In summary, we propose that the locomotion and postures of the S + L + P agent enhance social presence and engagement, as indicated by subjective questionnaires. However, these behaviors may divert participants' attention from the game, as evidenced by a lower EEG engagement index and supported by behavioral results.

6.3. Implications for human-agent interaction design

As noted by Oertel et al. (2020), engagement in human-agent interaction can manifest in two forms: Social engagement, which occurs between the human user and the agent, and task engagement, which occurs between the human user and the task at hand. Corrigan et al. (2013) further extended the concept of social engagement and task engagement to social-task engagement which happens in scenarios where human users and agents work collaboratively on an explicit task.

In our study, participants played the Gobang game against the virtual agent, where users engaged with both the game logic and the virtual agent behaviors. In other words, both task engagement (i.e., engagement with the Gobang game logic) and social engagement (i.e., engagement with the virtual agent) were involved in our study. The EEG engagement index could potentially measure the overall social-task engagement because participants' cognitive activities were driven by the brain, which the EEG signals could capture. In contrast, the engagement questionnaire used by our study mainly focuses on social engagement because each statement in the questionnaire is related to the interaction with the virtual agent. Our behavioral measurements, like GD proportion, player average speed, rounds of play, and task completion time, could be related to task engagement. Therefore, in our study, the engagement questionnaire measures social engagement, behavioral metrics assess task engagement, and the EEG-based engagement index reflects overall engagement. As discussed in Section 6.1 and Section 6.2, the S + L + P agent in our study enhanced social perceptions but may have led to lower overall engagement. There was no significant difference in task engagement as measured by behavioral metrics. However, to comprehensively understand the relationships among social engagement, task engagement, and overall social-task engagement in human-agent interactions, further research is necessary.

Additionally, responses to interview question Q4 on agent role preference revealed that participants with higher self-reported Gobang skill levels preferred the agent to take on an opponent role, while those with lower skill levels favored the assistant role. This suggests that users' individual preferences for the agent's role may also influence their perception of and interaction with the agent.

By integrating the discussions from this section with those in Section 6.1 and Section 6.2, we highlight the Design Implications (DI) of our work for future human-agent social interaction in task-oriented scenarios within MR:

DI1. Integrating both locomotion and postures can enhance the social perception of a MiRA (based on discussions in Section 6.1).

DI2. Although locomotion and postures in MiRAs can enhance social perception, they may also distract users, potentially leading to decreased task performance (based on discussions in Section 6.2).

DI3. A comprehensive evaluation of users' experiences with MiRAs should incorporate subjective, behavioral, and physiological measurements to capture a holistic view of user perception (based on discussions in Section 6.1 and Section 6.2).

DI4. Effective human-agent social interaction design should carefully balance social engagement, task engagement, and overall social-task integration to prevent enhanced social perception from leading to decreased overall engagement (based on discussions in Section 6.3).

DI5. The roles and behaviors of virtual agents should take into account users' individual preferences to optimize interaction experiences (based on observations in interview Q4).

7. Limitations

Our work has some limitations. First, this work did not directly assess task engagement and did not extensively investigate the interplay among social engagement, task engagement, and overall engagement. Second, although the S + L and S + L + P agents embodied in the MR environment, the locomotion behavior didn't show awareness of the physical environment, which was proved to be useful to improve social presence and engagement (Kim et al., 2017; 2018a). Moreover, although we reduced the transparency of the virtual agent's body to avoid occluding the chessboard behind the body, the realism of the virtual agent's appearance was also weakened. Another issue was that mounting the Hololens headset on top of the EEG cap might disturb the movements of electrodes and thus result in poor EEG signals.

8. Conclusion and future work

Nonverbal cues in MiRAs are essential for shaping users' perceptions, as reflected by physiological states captured through EEG. However, previous studies have primarily assessed the impact of MiRAs' nonverbal behaviors on social perception using subjective and behavioral measurements. In this paper, we investigate how nonverbal cues, such as locomotion and emotional postures in MiRAs, influence social perceptions by integrating EEG with behavioral and subjective measurements to assess users' social perceptions comprehensively.

In this paper, we explored the impact of mixed reality agents' non-verbal cues, like locomotion and emotional postures, on perceived social presence, engagement, and user brain activity. We presented a within-subject user study where participants were asked to win against the virtual agent over a Gobang game in an MR environment. Three types of the virtual agent were designed by varying the embodiment and behaviors: 1) A **S**peech-only (**S**) agent communicating through speech only, 2) an embodied agent with both **S**peech and **L**ocomotion behavior (**S + L**) for placing the pieces, and 3) an embodied agent with **S**peech, **L**ocomotion and **P**ostures (**S + L + P**) related to the game process. We found that the **S + L + P** agent had significantly higher social presence, engagement, and emotional arousal than the **S** agent and was liked most by the participants. However, the EEG-based engagement index at AF3 was lower in the **S + L + P** condition than in the **S + L** condition, opposite to the engagement measured by the questionnaire. No significant difference was found in behavioral measurements. We discussed the observed discrepancy between engagement questionnaire results and EEG-based engagement index and provided design implications for future mixed reality agent design.

In the future, we plan to explore the impact of relationships between the user and virtual agent on user perception and behaviors. For example, unlike playing against the virtual opponent shown in this paper, we can also have a virtual assistant to suggest where to place the piece while playing the Gobang game. It would also be interesting to explore how non-verbal cues help establish rapport, trust, and perceived friendliness when the virtual agent's role turns to an assistant.

Notes

1. https://developer.amazon.com/en-US/alexa
2. https://www.microsoft.com/en-us/cortana/
3. https://www.apple.com/siri/
4. https://www.soulmachines.com/
5. https://cloud.ibm.com/catalog/services/text-to-speech
6. https://www.mixamo.com/
7. https://dotween.demigiant.com/
8. https://developer.oculus.com/documentation/unity/audio-ovrlipsync-viseme-reference/
9. https://github.com/SSSxCCC/AlphaZero-In-Unity
10. https://github.com/microsoft/MixedRealityToolkit-Unity
11. https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/features/tools/holographic-remoting?view=mrtkunity-2022-05
12. https://www.unicorn-bi.com/
13. https://www.gtec.at/product/gnautilus-pro/
14. https://www.youtube.com/watch?v=LOfIr2F7-Tc
15. https://github.com/sccn/liblsl
16. https://mne.tools/stable/generated/mne.baseline.rescale.html

Disclosure statement

No potential conflict of interest was reported by the author(s).

Ethical approval

The studies involving human participants were reviewed and approved by the University of Auckland Human Participants Ethics Committee (UAHPEC). The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Funding

This work was supported by the China Scholarship Council under Grant CSC NO. 202008320270.

ORCID

Zhuang Chang http://orcid.org/0000-0003-0972-0752
Jiashuo Cao http://orcid.org/0000-0003-0779-7692
Kunal Gupta http://orcid.org/0000-0003-3963-8856
Huidong Bai http://orcid.org/0000-0001-8252-1122
Mark Billinghurst http://orcid.org/0000-0003-4172-6759

References

Alimardani, M., van den Braak, S., Jouen, A.-L., Matsunaka, R., & Hiraki, K. (2021). *Assessment of engagement and learning during child-robot interaction using EEG signals* [Paper presentation]. In Social Robotics: 13th International Conference, ICSR 2021, Proceedings 13 (pp. 671–682). https://doi.org/10.1007/978-3-030-90525-5_59

Allameh, M., & Zaman, L. (2021). *Jessy: A conversational assistant for tutoring digital board games* [Paper presentation]. Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play (pp. 168–173). https://doi.org/10.1145/3450337.3483496

Antonenko, P., Paas, F., Grabner, R., & Van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review*, *22*(4), 425–438. https://doi.org/10.1007/s10648-010-9130-y

Antonio Gómez Jáuregui, D., Giraud, T., Isableu, B., & Martin, J. (2021). Design and evaluation of postural interactions between users and a listening virtual agent during a simulated job interview. *Computer Animation and Virtual Worlds*, *32*(6), e2029. https://doi.org/10.1002/cav.2029

Barambones, J., Cano-Benito, J., Sánchez-Rivero, I., Imbert, R., & Richoux, F. (2023). Multiagent systems on virtual games: A systematic mapping study. *IEEE Transactions on Games*, *15*(2), 134–147. https://doi.org/10.1109/TG.2022.3214154

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., Olmstead, R. E., Tremoulet, P. D., & Craven, P. L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine*, *78*(5 Suppl), B231–B244.

Bocchi, A., Palmiero, M., Persichetti, F., Matteoli, M., Guariglia, C., & Piccardi, L. (2024). Giant chess game enhances spatial navigational skills in 6-years-old children: Preliminary findings. *Applied Neuropsychology, Child*, *13*(1), 37–44. https://doi.org/10.1080/21622965.2022.2112680

Busch, N. A., & VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(37), 16048–16053. https://doi.org/10.1073/pnas.1004801107

Castiblanco Jimenez, I. A., Gomez Acevedo, J. S., Olivetti, E. C., Marcolin, F., Ulrich, L., Moos, S., & Vezzetti, E. (2022). User engagement comparison between advergames and traditional advertising using eeg: Does the user's engagement influence purchase intention? *Electronics*, *12*(1), 122. https://doi.org/10.3390/electronics12010122

Chang, Z., Bai, H., Zhang, L., Gupta, K., He, W., & Billinghurst, M. (2022). The impact of virtual agents' multimodal communication on brain activity and cognitive load in virtual reality. *Frontiers in Virtual Reality*, *3*, 179. https://doi.org/10.3389/frvir.2022.995090

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, *62*(1), 73–101. https://doi.org/10.1146/annurev.psych.093008.100427

Coelli, S., Sclocco, R., Barbieri, R., Reni, G., Zucca, C., & Bianchi, A. M. (2015). *Eeg-based index for engagement level monitoring during sustained attention* [Paper presentation]. In 2015 37th Annual International Conference of the Ieee Engineering in Medicine and Biology Society (EMBC) (pp. 1512–1515). https://doi.org/10.1109/EMBC.2015.7318658

Corrigan, L. J., Peters, C., Castellano, G., Papadopoulos, F., Jones, A., Bhargava, S., & Aylett, R. (2013). Social-task engagement: Striking a balance between the robot and the task. *Embodied Communication, Goals Intentions Workshop Icsr*, *13*, 1–7.

Damette, N., Szymanski, M., Mualla, Y., Tchappi, I., Najjar, A., & Adda, M. (2024). Forecasting future behavior: Agents in board game strategy. *Procedia Computer Science*, *241*, 187–194. https://doi.org/10.1016/j.procs.2024.08.026

Dehais, F., Lafont, A., Roy, R., & Fairclough, S. (2020). A neuroergonomics approach to mental workload, engagement and human performance. *Frontiers in Neuroscience*, *14*, 268. https://doi.org/10.3389/fnins.2020.00268

DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., & Gainer, A, et al. (2014). *Simsensei kiosk: A virtual human interviewer for healthcare decision support* [Paper presentation]. Proceedings of the 2014 international conference on autonomous agents and multi-agent systems (pp. 1061–1068). https://doi.org/10.5555/2615731.2617415

Eichhorn, C., Plecher, D. A., Golovnya, O., Volkert, D., Hiyama, A., & Klinker, G. (2021). *Tangible chess for dementia patients–playing with conductive 3d printed figures on a touchscreen* [Paper presentation]. International conference on human-computer interaction (pp. 38–57). https://doi.org/10.1007/978-3-030-78111-8_3

Freeman, F. G., Mikulka, P. J., Prinzel, L. J., & Scerbo, M. W. (1999). Evaluation of an adaptive automation system using three eeg indices with a visual tracking task. *Biological Psychology*, *50*(1), 61–76. https://doi.org/10.1016/s0301-0511(99)00002-2

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*(267), 267. https://doi.org/10.3389/fnins.2013.00267

Gregory, S. E., Kelly, C. L., & Kessler, K. (2021). Look into my "virtual" eyes: What dynamic virtual agents add to the realistic study of joint attention. *Frontiers in Virtual Reality*, *2*, 798899. https://doi.org/10.3389/frvir.2021.798899

Gupta, K., Hajika, R., Pai, Y. S., Duenser, A., Lochner, M., & Billinghurst, M. (2020). *Measuring human trust in a virtual assistant using physiological sensing in virtual reality* [Paper presentation]. 2020 IEEE Conference on Virtual Reality and 3d User Interfaces (vr) (pp. 756–765). https://doi.org/10.1109/VR46266.2020.00099

Harms, C., & Biocca, F. (2004). Internal consistency and reliability of the networked minds measure of social presence. In *Seventh annual international workshop: Presence.* (Vol. 2004). Universidad Politécnica de Valencia.

Hart, S. G., & Staveland, L. E. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139–183). Elsevier. https://doi.org/10.1016/S0166-4115(08)62386-9

Holz, T., Campbell, A. G., O'Hare, G. M., Stafford, J. W., Martin, A., & Dragone, M. (2011). Mira—mixed reality agents. *International Journal of Human-Computer Studies*, *69*(4), 251–268. https://doi.org/10.1016/j.ijhcs.2010.10.001

Huang, A., Knierim, P., Chiossi, F., Chuang, L. L., & Welsch, R. (2022). *Proxemics for human-agent interaction in augmented reality* [Paper presentation]. In Proceedings of the 2022 Chi Conference on Human Factors in Computing Systems (pp. 1–13). https://doi.org/10.1145/3491102.3517593

Ishii, R., Ahuja, C., Nakano, Y. I., & Morency, L.-P. (2020). *Impact of personality on nonverbal behavior generation. In Proceedings of the 20th* [Paper presentation]. ACM International Conference on Intelligent Virtual Agents (pp. 1–8). https://doi.org/10.1145/3383652.3423908

Jung, Y., Jung, G., Jeong, S., Kim, C., Woo, W., Hong, H., & Lee, U. (2023). "enjoy, but moderately!": Designing a social companion robot for social engagement and behavior moderation in solitary drinking context. *Proceedings of the ACM on Human-Computer Interaction*, *7*(CSCW2), 1–24. https://doi.org/10.1145/3610028

Katsarou, E., Wild, F., Sougari, A.-M., & Chatzipanagiotou, P. (2023). A systematic review of voice-based intelligent virtual agents in efl education. *International Journal of Emerging Technologies in Learning (iJET)*, *18*(10), 65–85. https://doi.org/10.3991/ijet.v18i10.37723

Kevin, S., Pai, Y. S., & Kunze, K. (2018). *Virtual gaze: Exploring use of gaze as rich interaction method with virtual agent in interactive virtual reality content* [Paper presentation]. Proceedings of the 24th Acm Symposium on Virtual Reality Software and Technology (pp. 1–2). https://doi.org/10.1145/3281505.3281587

Kim, K., Boelling, L., Haesler, S., Bailenson, J., Bruder, G., & Welch, G. F. (2018a). *Does a digital assistant need a body? the influence of visual embodiment and social behavior on the perception of intelligent virtual agents in AR* [Paper presentation]. 2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (pp. 105–114). https://doi.org/10.1109/ISMAR.2018.00039

Kim, K., Boelling, L., Haesler, S., Bailenson, J., Bruder, G., & Welch, G. F. (2018b). *Does a digital assistant need a body? the influence of visual embodiment and social behavior on the perception of intelligent virtual agents in AR* [Paper presentation]. 2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (pp. 105–114). https://doi.org/10.1109/ISMAR.2018.00039

Kim, K., de Melo, C. M., Norouzi, N., Bruder, G., & Welch, G. F. (2020). *Reducing task load with an embodied intelligent virtual assistant for improved performance in collaborative decision making* [Paper presentation]. 2020 IEEE Conference on Virtual Reality and 3d User Interfaces (vr), (pp. 529–538). https://doi.org/10.1109/VR46266.2020.00074

Kim, K., Maloney, D., Bruder, G., Bailenson, J. N., & Welch, G. F. (2017). The effects of virtual human's spatial and behavioral coherence with physical objects on social presence in ar. *Computer Animation and Virtual Worlds*, *28*(3-4), e1771. https://doi.org/10.1002/cav.1771

Kruse, L., Hertel, J., Mostajeran, F., Schmidt, S., & Steinicke, F. (2023). *Would you go to a virtual doctor? a systematic literature review on user preferences for embodied virtual agents in healthcare* [Paper presentation]. 2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (pp. 672–682). https://doi.org/10.1109/ISMAR59233.2023.00082

Kum, J., & Lee, M. (2022). Can gestural filler reduce user-perceived latency in conversation with digital humans? *Applied Sciences*, *12*(21), 10972. https://doi.org/10.3390/app122110972

Kumar, N., & Kumar, J. (2016). Measurement of cognitive load in hci systems using eeg power spectrum: An experimental study. *Procedia Computer Science*, *84*, 70–78. https://doi.org/10.1016/j.procs.2016.04.068

Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *The American Psychologist*, *50*(5), 372–385. https://doi.org/10.1037/0003-066x.50.5.372

Lee, M., Bruder, G., Höllerer, T., & Welch, G. (2018). Effects of unaugmented periphery and vibrotactile feedback on proxemics with virtual humans in ar. *IEEE Transactions on Visualization and Computer Graphics*, *24*(4), 1525–1534. https://doi.org/10.1109/TVCG.2018.2794074

Lee, M., Norouzi, N., Bruder, G., Wisniewski, P. J., & Welch, G. F. (2018). *The physical-virtual table: Exploring the effects of a virtual human's physical influence on social interaction* [Paper presentation]. Proceedings of the 24th Acm Symposium on Virtual Reality Software and Technology (pp. 1–11). https://doi.org/10.1145/3281505.3281533

Lee, M., Norouzi, N., Bruder, G., Wisniewski, P. J., & Welch, G. F. (2021). Mixed reality tabletop gameplay: Social interaction with a virtual human capable of physical influence. *IEEE Transactions on Visualization and Computer Graphics*, *27*(8), 3534–3545. https://doi.org/10.1109/TVCG.2019.2959575

Leite, I., Castellano, G., Pereira, A., Martinho, C., & Paiva, A. (2014). Empathic robots for long-term interaction: Evaluating social presence, engagement and perceived support in children. *International Journal of Social Robotics*, *6*(3), 329–341. https://doi.org/10.1007/s12369-014-0227-1

Leite, I., Martinho, C., Pereira, A., & Paiva, A. (2009). *As time goes by: Long-term evaluation of social presence in robotic companions* [Paper presentation]. Ro-Man 2009-the 18th IEEE International Symposium on Robot and Human Interactive Communication (pp. 669–674). https://doi.org/10.1109/ROMAN.2009.5326256

Li, C., Androulakaki, T., Gao, A. Y., Yang, F., Saikia, H., Peters, C., & Skantze, G. (2018). *Effects of posture and embodiment on social distance in human-agent interaction in mixed reality* [Paper presentation]. Proceedings of the 18th International Conference on Intelligent Virtual Agents (pp. 191–196). https://doi.org/10.1145/3267851.3267870

Li, X., Zhang, W., Chen, J., Wu., & L., Cairangdanghzhou. (2022). *Cognitive Systems and Information Processing: 6th International Conference, Iccsip 2021, Gobang game algorithm based on reinforcement learning. In, suzhou, china, november 20–21, 2021, revised selected papers 6* (pp. 463–475). https://doi.org/10.1007/978-981-16-9247-5_36

Liu, H., Choi, M., Kao, D., & Mousas, C. (2023). Synthesizing game levels for collaborative gameplay in a shared virtual environment. *ACM Transactions on Interactive Intelligent Systems*, *13*(1), 1–36. https://doi.org/10.1145/3558773

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., & Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(18), 8135–8139. https://doi.org/10.1073/pnas.92.18.8135

McMahan, T., Parberry, I., & Parsons, T. D. (2015). Evaluating player task engagement and arousal using electroencephalography. *Procedia Manufacturing*, *3*, 2303–2310. https://doi.org/10.1016/j.promfg.2015.07.376

Mendes, D. C., Costa, R., Agrela, G., I Badia, S. B., & Cameirão, M. S. (2024). *Evaluating the impact of anthropomorphism and verbal interaction on users' perception of a virtual agent in facilitating self-disclosure* [Paper presentation]. 2024 IEEE 12th International Conference on Serious Games and Applications for Health (SEGAH) (pp. 1–8). https://doi.org/10.1109/SeGAH61285.2024.10639572

Milcent, A.-S., Kadri, A., & Richir, S. (2022). Using facial expressiveness of a virtual agent to induce empathy in users. *International Journal of Human–Computer Interaction*, *38*(3), 240–252. https://doi.org/10.1080/10447318.2021.1938387

Miller, M. R., Jun, H., Herrera, F., Yu Villa, J., Welch, G., & Bailenson, J. N. (2019). Social interaction in augmented reality. *PloS One*, *14*(5), e0216290. https://doi.org/10.1371/journal.pone.0216290

Mouton, A., Gillet, N., Mouton, F., Van Kann, D., Bruyère, O., Cloes, M., & Buckinx, F. (2017). Effects of a giant exercising board game intervention on ambulatory physical activity among nursing home residents: A preliminary study. *Clinical Interventions in Aging*, *12*, 847–858. https://doi.org/10.2147/CIA.S134760

Nijholt, A. (2022). *Towards social companions in augmented reality: Vision and challenges* [Paper presentation]. International Conference on Human-Computer Interaction (pp. 304–319). https://doi.org/10.1007/978-3-031-05431-0_21

Norouzi, N., Bruder, G., Belna, B., Mutter, S., Turgut, D., & Welch, G. (2019). A systematic review of the convergence of augmented reality, intelligent virtual agents, and the internet of things. In *Artificial Intelligence in IoT* (pp. 1–24). https://doi.org/10.1007/978-3-030-04110-6_1

Norouzi, N., Kim, K., Bruder, G., Erickson, A., Choudhary, Z., Li, Y., & Welch, G. (2020). *A systematic literature review of embodied augmented reality agents in head-mounted display environments* [Paper presentation]. Proceedings of the International Conference on Artificial Reality and Telexistence & Eurographics Symposium on Virtual Environments.

Nuamah, J., & Seong, Y. (2018). Support vector machine (SVM) classification of cognitive tasks based on electroencephalography (EEG) engagement index. *Brain-Computer Interfaces*, *5*(1), 1–12. https://doi.org/10.1080/2326263X.2017.1338012

Oertel, C., Castellano, G., Chetouani, M., Nasir, J., Obaid, M., Pelachaud, C., & Peters, C. (2020). Engagement in human-agent interaction: An overview. *Frontiers in Robotics and AI*, *7*, 92. https://doi.org/10.3389/frobt.2020.00092

Piette, E., Stephenson, M., Soemers, D. J., & Browne, C. (2021). *General board game concepts* [Paper presentation]. 2021 IEEE Conference on Games (Cog) (pp. 01–08). https://doi.org/10.1109/CoG52621.2021.9618990

Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology*, *40*(1-2), 187–195. https://doi.org/10.1016/0301-0511(95)05116-3

Pretty, E. J., Fayek, H. M., & Zambetta, F. (2024). A case for personalized non-player character companion design. *International Journal of Human–Computer Interaction*, *40*(12), 3051–3070. https://doi.org/10.1080/10447318.2023.2181125

Rajendran, V., Jayalalitha, S., & Adalarasu, K. (2022). Eeg based evaluation of examination stress and test anxiety among college students. *IRBM*, *43*(5), 349–361. https://doi.org/10.1016/j.irbm.2021.06.011

Randhavane, T., Bera, A., Kapsaskis, K., Gray, K., & Manocha, D. (2019). Fva: Modeling perceived friendliness of virtual agents using movement characteristics. *IEEE Transactions on Visualization and Computer Graphics*, *25*(11), 3135–3145. https://doi.org/10.1109/TVCG.2019.2932235

Riedmann, A., Schaper, P., & Lugrin, B. (2024). Integration of a social robot and gamification in adult learning and effects on motivation, engagement and performance. *AI & Society*, *39*(1), 369–388. https://doi.org/10.1007/s00146-022-01514-y

Ruqeyya, G., Hafeez, T., Saeed, S. M. U., & Ishwal, A. (2022). *EEG-based engagement index for video game players* [Paper presentation]. 2022 International Conference on Emerging Trends in Electrical, Control, and Telecommunication Engineering (ETECTE) (pp. 1–6). https://doi.org/10.1109/ETECTE55893.2022.10007386

Sidner, C. L., Kidd, C. D., Lee, C., & Lesh, N. (2004). *Where to look: A study of human-robot engagement* [Paper presentation]. In Proceedings of the 9th International Conference on Intelligent User Interfaces (pp. 78–84). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/964442.964458

Sun, N., Botev, J., Khaluf, Y., & Simoens, P. (2022). *Theory of mind and delegation to robotic virtual agents* [Paper presentation]. 2022 31st IEEE International Conference on Robot and Human Interactive Communication (ro-Man) (pp. 454–460). https://doi.org/10.1109/RO-MAN53752.2022.9900789

Szafir, D., & Mutlu, B. (2012). *Pay attention! designing adaptive agents that monitor and improve user engagement* [Paper presentation]. Proceedings of the Sigchi Conference on Human Factors in Computing Systems (pp. 11–20). https://doi.org/10.1145/2207676.2207679

Torre, R., Fua, P., Balcisoy, S., Ponder, M., & Thalmann, D. (2000). *Interaction between real and virtual humans: Playing checkers* [Paper presentation]. Virtual Environments 2000: Proceedings of the Eurographics Workshop in Amsterdam, The Netherlands (pp. 23–32). https://doi.org/10.1007/978-3-7091-6785-4_4

Vrins, A., Pruss, E., Prinsen, J., Ceccato, C., & Alimardani, M. (2022). *Are you paying attention? the effect of embodied interaction with an adaptive robot tutor on user engagement and learning performance* [Paper presentation]. International Conference on Social Robotics (pp. 135–145). https://doi.org/10.1007/978-3-031-24670-8_13

Wang, H., Preuss, M., & Plaat, A. (2021). *Adaptive warm-start mcts in alphazero-like deep reinforcement learning* [Paper presentation]. Pricai 2021: Trends in Artificial Intelligence: 18th Pacific Rim International Conference on Artificial Intelligence, Pricai 2021, Hanoi, Vietnam, November 8–12, 2021, proceedings, part iii (pp. 60–71). https://doi.org/10.1007/978-3-030-89370-5_5

Wang, I., & Ruiz, J. (2021). Examining the use of nonverbal communication in virtual agents. *International Journal of Human–Computer Interaction*, *37*(17) 1–26. https://doi.org/10.1080/10447318.2021.1898851

Wang, I., Smith, J., & Ruiz, J. (2019). *Exploring virtual agents for augmented reality* [Paper presentation]. Proceedings of the 2019 chi conference on human factors in computing systems (pp. 1–12). https://doi.org/10.1145/3290605.3300511

Welch, P. (1967). The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, *15*(2), 70–73. https://doi.org/10.1109/TAU.1967.1161901

Ye, Z.-M., Chen, J.-L., Wang, M., & Yang, Y.-L. (2021). *Paval: Position-aware virtual agent locomotion for assisted virtual reality navigation* [Paper presentation]. 2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (pp. 239–247). https://doi.org/10.1109/ISMAR52148.2021.00039

Yousefi, M., Crowe, S. E., Hoermann, S., Sharifi, M., Romera, A., Shahi, A., & Piumsomboon, T. (2024). Advancing prosociality in extended reality: Systematic review of the use of embodied virtual agents to trigger prosocial behaviour in extended reality. *Frontiers in Virtual Reality*, *5*, 1386460. https://doi.org/10.3389/frvir.2024.1386460

About the authors

Zhuang Chang is a PhD student in the Empathic Computing Laboratory at the Auckland Bioengineering Institute, University of Auckland. His research focuses on Empathic Mixed Reality Agents, investigating empathy in virtual agents within mixed-reality environments using physiological sensors and tactile actuators to enhance human-agent interactions.

Jiashuo Cao is a PhD student at the Empathic Computing Laboratory, University of Auckland. His research interests covered HCI, XR, and Psychotherapy. He collaborates closely with psychotherapists to explore how virtual agents and virtual reality can enhance remote psychotherapy.

Kunal Gupta is a Research Fellow at the Empathic Computing Laboratory, University of Auckland. His research bridges human-AI interaction and empathic computing, using physiological signals to create adaptive systems that enhance mental well-being and support cognitive rehabilitation, with applications in both virtual environments and real-world health tech interventions.

Huidong Bai is a Research Fellow at the Empathic Computing Laboratory, University of Auckland. His research focuses on developing Mixed Reality interfaces for remote collaboration, emphasizing general enhancements in spatial understanding and empathic computing to improve mutual communications.

Mark Billinghurst directs the Empathic Computing Laboratory at the University of Auckland and University of South Australia. He researches how virtual and real worlds can be merged, publishing over 750 papers on AR, VR, collaboration, Empathic Computing, and related topics. In 2023 he was elected as an IEEE Fellow.