International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: www.tandfonline.com/journals/hihc20

Postures and Locomotion in Mixed Reality Agents: Effects on Social Perception of Virtual Opponents and Assistants

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

To cite this article: Zhuang Chang, Kunal Gupta, Jiashuo Cao, Huidong Bai & Mark Billinghurst (17 Nov 2025): Postures and Locomotion in Mixed Reality Agents: Effects on Social Perception of Virtual Opponents and Assistants, International Journal of Human–Computer Interaction, DOI: 10.1080/10447318.2025.2583473

To link to this article: https://doi.org/10.1080/10447318.2025.2583473

Published online: 17 Nov 2025.

Submit your article to this journal

Article views: 33

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=hihc20

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2025.2583473

Taylor & Francis
Taylor & Francis Group

Check for updates

Postures and Locomotion in Mixed Reality Agents: Effects on Social Perception of Virtual Opponents and Assistants

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

Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand

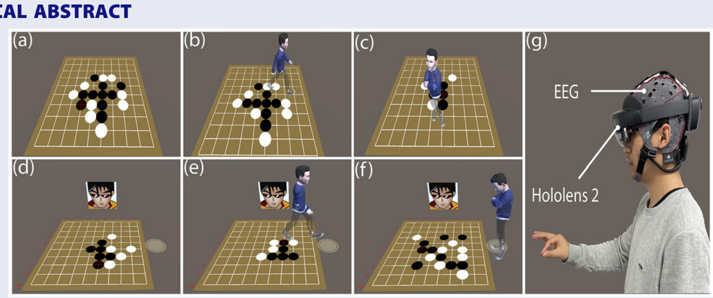
ABSTRACT

Integrating non-verbal cues into Mixed Reality Agents (MiRAs) enhances their ability to engage users and foster socially rich interactions. This paper investigates the role of locomotion and body posture in shaping user engagement, social presence, and interaction quality through two user studies involving a turn-based Gobang game. From these studies we found that in a competitive context, MiRAs' locomotion and posture enhanced social presence and engagement, but while in a cooperative context, these behaviors fostered rapport but not trust. By integrating subjective, behavioral, and physiological measures, including EEG, this study provides a holistic understanding of MiRAs' impact. The findings offer actionable design implications for creating engaging and socially effective virtual agents, advancing the field of human-agent interaction in Mixed Reality. Future research directions include exploring long-term effects, using diverse application domains, and supporting multimodal interactions.

KEYWORDS

Mixed reality agents; social perception; opponent; assistant; EEG

GRAPHICAL ABSTRACT



HIGHLIGHTS

- Postures and Locomotion are key non-verbal cues influencing user engagement and social perceptions of Mixed Reality Agents (MiRAs).
- MiRAs' posture and locomotion can enhance social presence and engagement in competitive contexts and foster rapport, but not trust, in cooperative contexts.
- An EEG-based engagement index provides deeper insights into the cognitive effects of MiRAs' non-verbal behaviors compared to traditional subjective and behavioral measures.

1. Introduction

This paper explores how the body posture and Locomotion of Mixed Reality Agents can affect how they are perceived. Integrating Intelligent Virtual Agents (IVAs) into various applications, such as virtual assistants, educational tools, healthcare support systems, and entertainment platforms, has significantly transformed the landscape of Human-Computer Interaction (HCI). Within the broader ecosystem of IVAs, researchers typically identify three main categories: 1) Voice-only agents (Katsarou et al., 2023) that exist without any visual form (e.g., Amazon Alexa,[1] Microsoft Cortana,[2] and Apple

CONTACT Zhuang Chang zcha621@aucklanduni.ac.nz Auckland Bioengineering Institute, University of Auckland, 70 Symonds Street, Auckland 1010, New Zealand
© 2025 Taylor & Francis Group, LLC

Siri[3]), 2) embodied agents (Yousefi et al., 2024) possessing a physical or virtual body, such as Soul Machines' conversational virtual characters[4] or Non-Player Characters (NPCs) in computer games (Pretty et al., 2024), and 3) abstract or partially embodied agents (Gupta et al., 2024), which use simplified or stylized representations (e.g., an aura, blob, or floating orb) instead of humanoid forms. Despite the commercial success of voice-only agents, previous research has demonstrated that agents capable of using non-verbal cues often provide more engaging and socially rich interactions for users (Mendes et al., 2024; Kruse et al., 2023).

In this work, we focus on embodied agents integrated into Mixed Reality (MR) environments, referred to as Mixed Reality Agents (MiRAs) (Holz et al., 2011). Unlike voice-only agents, MiRAs incorporate non-verbal cues such as facial expressions, gaze, body posture, and locomotion to communicate with users (Norouzi et al., 2020; Wang & Ruiz, 2021). These non-verbal behaviors, essential for natural human-human communication, significantly shape users' perception of the agents. Specifically, locomotion has been shown to enhance the agent's naturalness and believability by enabling it to move within the user's physical space (Kim et al., 2018), while body posture influences users' willingness to engage with the agent and their perception of its social presence (Li et al., 2018).

Despite the growing body of research exploring MiRAs, most studies have primarily relied on subjective (e.g., questionnaires) and behavioral (e.g., task performance) measures. While these approaches provide valuable insights, physiological measures such as Electroencephalography (EEG) offer a unique opportunity to gain a deeper understanding of users' cognitive and emotional responses during the experimental task rather than relying on post-task evaluations. EEG has been successfully used to assess cognitive load (Kumar & Kumar, 2016), trust in virtual assistants (Gupta et al., 2020), and engagement in gaming (Ruqeyya et al., 2022; Pope et al., 1995). However, the role of non-verbal cues, such as locomotion and posture, in influencing user engagement with MiRAs as measured by EEG remains relatively underexplored.

To address this gap, we conducted two user studies to explore the impact of MiRAs' locomotion and body posture on users' engagement, social presence, and interaction quality in MR. Both studies involved a turn-based Gobang game, chosen for its strategic depth and suitability for goal-oriented interaction. The first study, previously published in Chang et al. (2024), investigated the effects of MiRAs' locomotion and posture in a competitive context where the MiRA acted as an *opponent*. The second study builds upon these findings by examining the MiRA in a cooperative context, where it functioned as an *assistant*, providing strategic hints on-demand to help the user outperform a computer opponent. This dual-context approach allowed us to investigate the role-dependent effects of MiRAs' non-verbal cues across varying interaction dynamics.

The primary research question addressed in this paper is: *How do the locomotion and posture of Mixed Reality Agents influence user engagement and social perceptions in competitive and cooperative contexts?*

Our work contributes in three ways:

C1. We investigate the impact of MiRAs' non-verbal cues on users' social perceptions in competitive and cooperative contexts by integrating EEG data with subjective and behavioral measures.

C2. Building on our previously published study featuring a virtual opponent (Chang et al., 2024), this paper presents a new user study investigating how a virtual assistant's locomotion and postures influence rapport building.

C3. We offer novel design implications for human-agent interaction in Mixed Reality (MR), derived from comparing two user studies featuring virtual agents as opponents or assistants.

The remainder of this paper is organized as follows: Section 2 reviews related work, situating our research within the broader context and highlighting its novelty. Section 3 describes the system used for both studies, providing technical details of the experimental setup. Section 4 presents an overview of the two studies, outlining their objectives and methodologies. Section 5 summarizes the key findings from Study 1, while Section 6 delves into the details of Study 2. Section 7 discusses the design implications derived from both studies and addresses their limitations. Finally, Section 8 concludes the paper and suggests directions for future research.

2. Related work

In this study, we investigate how MiRAs' locomotion and body postures affect social perception within a board game context using EEG as a physiological measure of engagement. To establish the foundation for our research, we review related work in four key areas: 1) Non-verbal cues in embodied virtual agents, 2) social perceptions of embodied virtual agents, 3) virtual agents in board games, and 4) EEG for measuring engagement and attention.

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), play a key role in human-agent social interactions (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. The design incorporated non-verbal cues like facial expressions, gaze, gestures, and postures to establish a strong therapist-user relationship. Similarly, He et al. (2022) assessed data-driven co-speech gestures of embodied conversational agents and found that data-driven generated gestures attracted more participants' attention. These studies highlight the vital role of non-verbal cues in fostering meaningful interactions between users and embodied virtual agents.

Research has shown that 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) investigated the impact of agent postures (closed vs. open postures) and embodiment (robot vs. human and virtual vs. physical) on social distance in the MR environment. Their findings on agent postures showed that users maintained greater social distance from agents displaying open postures, yet were more engaging and willing to interact with them. Similarly, Randhavane et al. (2019a) presented a data-driven friendliness model-based algorithm to control the virtual agents' gaits, gestures, and gaze behaviors in Augmented Reality (AR). In their user study, they found that the embodied virtual agent driven by their friendliness model was perceived to have higher friendliness and social presence. In addition to virtual agents embodied in AR, virtual agents embodied in immersive Virtual Reality (VR) that display postures were shown to increase co-presence and capture more attention, as evidenced by EEG measurements (Chang et al., 2022). These studies demonstrate that the postures of virtual agents in 3D immersive extended reality (XR) environments (i.e., VR/AR/MR environments) 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).

Alongside posture, locomotion is a critical aspect of human-agent interaction, particularly for virtual agents embodied in XR environments (Holz et al., 2011; Nijholt, 2022; Norouzi et al., 2019). For instance, Kim et al. (2018) conducted a user study in which participants interacted with a virtual agent to complete tasks within a lab environment, such as switching off a real-world light bulb, checking the lab temperature, contacting other lab members, and leaving a private space. The study compared three agent designs: (1) a disembodied speech-only agent, (2) an embodied agent with speech and gestures, and (3) an embodied agent with speech, gestures, and locomotion in AR space. The findings showed that incorporating embodied gestures and locomotion that demonstrated awareness of and interaction with the surrounding environment significantly enhanced users' trust in the agent and their perception of its social presence. Similarly, Kim et al. (2017) demonstrated that a virtual human seated in a wheelchair, moving through an AR space while avoiding physical obstacles like a chair in the real world, enhanced social presence. This interaction also influenced users' behavior, prompting them to avoid walking through the virtual human space. Beyond research on locomotion in virtual agents embodied in AR space, Ye et al. (2021) focused on embodied virtual agents' adaptive locomotion in VR. They introduced a real-time position-aware virtual agent locomotion method based on users' positions and surrounding virtual objects for room-scale VR navigation. In addition to the awareness of the environment associated with locomotion in virtual agents embodied in VR/AR environments, locomotion involves changes in spatial positioning, so it can also affect proxemics, the physical distance users maintain from virtual agents during interactions (Huang et al., 2022; Lee et al., 2018a; Miller et al., 2019). In short, locomotion in embodied virtual agents affects users' social perceptions and behaviors.

While previous research has explored the effects of postures and locomotion in virtual agents embodied in XR, it has predominantly relied on subjective and behavioral measurements. In our work, we assess the effects of MiRAs' locomotion and postures by incorporating subjective, behavioral, and physiological measures (i.e., EEG). Locomotion is utilized to convey the virtual agent's intended spatial positions in the MR environment, while postures are designed to express the agent's emotional state. Through these designed locomotion and postures, our primary focus is on influencing user engagement during a Gobang game and evaluating the engagement using both questionnaires and EEG.

2.2. Social perceptions of embodied virtual agents

In human-human social interactions, understanding others' behaviors through observing and interpreting their actions is a fundamental aspect of daily communication (Grèzes & de Gelder, 2008). This process is closely linked to social perception, which involves recognizing and interpreting social cues, such as eye gaze, facial expressions, body movements, and other forms of biological motion, to evaluate social roles and relationships (Allison et al., 2000; McCleery et al., 2014). Similarly, in human-agent social interactions, social perception involves interpreting the agent's behaviors conveyed by verbal and non-verbal cues to infer its intentions, roles, and relational dynamics. These cues enable users to engage with embodied virtual agents as if they were social entities, fostering a sense of connection and enhancing the interaction's effectiveness and realism (Kim et al., 2018; Wang et al., 2019).

The social perception of embodied virtual agents is a broad concept, commonly evaluated through specific and interrelated components such as social presence (Tan & Liew, 2020), engagement (Oertel et al., 2020), rapport (Sun, 2023), trust (Bickmore & Cassell, 2001), likeability (Jeong et al., 2023), and emotional expressiveness (Aneja et al., 2021), among others. These components each contribute uniquely to the perception of virtual agents and are critical for understanding user interactions. Social Presence is defined as the degree to which a user perceives access to the intelligence, intentions, and sensory impressions of another (Biocca, 1997; Pereira et al., 2014), and is often measured using tools such as the Networked Minds Social Presence Questionnaire (Harms & Biocca, 2004). Engagement, described as "the process by which two (or more) participants establish, maintain, and end their perceived connection" (Sidner et al., 2004), is commonly assessed through self-report scales (Leite et al., 2014) or behavioral observation (Oertel et al., 2020). Trust, reflecting the user's willingness to rely on the agent based on its perceived reliability and competence (Bickmore & Cassell, 2001), can be evaluated using instruments like the Automated System Trust Scales Questionniare (Jian et al., 2000). Rapport, denoting a sense of harmony and mutual understanding between the user and the agent (Sun, 2023), can be measured through Human-Agent Rapport Questionnaire (Cerekovic et al., 2016). Finally, emotional arousal and valence, which capture the intensity and positivity/negativity of emotional experiences, are assessed using tools like the Self-Assessment Manikin (SAM) questionnaire (Bradley & Lang, 1994).

In addition to subjective questionnaires commonly used to measure components of social perception, behavioral indices, and physiological measures have also been employed to provide deeper insights. For instance, Andrist et al. (2017) demonstrated that bidirectional gaze between users and virtual characters enhanced engagement, as evidenced by reduced error rates and faster task completion times in a collaborative task. Similarly, Gupta et al. (2020) assessed users' trust in virtual agents during a shape selector task using a combination of physiological signals, including EEG, Galvanic Skin Response (GSR), and Heart Rate Variability (HRV), along with behavioral measures such as head movement and task performance.

However, limited attention has been given to evaluating the social perceptions of virtual agents embodied in MR by integrating physiological, behavioral, and subjective measurements. In this work, we present two user studies featuring virtual agents serving as an opponent and an assistant in a board game. We measured key social perception components, including social presence, engagement, emotional arousal, valence, and user preferences for the opponent. For the assistant, we additionally measured trust and rapport using validated questionnaires. Furthermore, we incorporated behavioral indices, such as task completion time and performance, alongside EEG data to comprehensively examine the impact of locomotion and postures on users' social perceptions of virtual agents in both competitive and cooperative contexts.

2.3. Virtual agents in board games

Integrating agents into multiplayer, turn-based board games has been largely explored in human-agent interaction research (Barambones et al., 2022; Damette et al., 2024; Piette et al., 2021). For example, Leite et al. (2014) designed an empathic model for a social robot that played chess with children on an electronic chessboard. This robot detected all moves made by both the player and itself, enabling it to understand the current game situation and determine appropriate empathic behaviors. Similarly, Sun et al. (2022) developed a Colored Trails board game in a VR environment to examine how a virtual opponent's theory of mind abilities influenced users' delegation behavior, requiring participants to negotiate with the agent for success. Beyond direct gameplay, Karim et al. (2023) investigated the role of virtual agents in improving rulebook accessibility and providing companionship in board games for individuals who are blind or have low vision. These structured, rule-based board game environments serve as ideal testbeds for exploring various aspects of human behavior in interaction with virtual agents.

In multiplayer, turn-based board games, virtual agents can act as opponents, competing against users, or as assistants, supporting users in achieving victory. For example, Eichhorn et al. (2021) created virtual chess opponents on a tablet screen to train players in strategic gameplay, while Allameh and Zaman (2021) developed a virtual assistant for The Royal Game of Ur, capable of answering rule-related questions, offering strategic move suggestions based on the player's state, and alerting players to important game events. The clear rules and structured nature of board games enable virtual agents to effectively monitor game flow, thereby enhancing user engagement.

Unlike screen-based virtual agents, those embodied in VR, AR, or MR environments have been integrated into board games to provide a more immersive and interactive experience (Lee et al., 2019; Liu et al., 2023; Torre et al., 2000). For example, Lee et al. (2018b) introduced a virtual human in AR seated at a physical table to play a tabletop game, where players take turns moving their tokens along designated spots on the shared surface. This setup enhances the sense of presence and enables more natural interactions between players and the virtual agent. However, despite these advances in integrating AR/MR virtual agents into board games, the limited Field of View (FoV) of AR/MR headsets like Hololens 2 often impairs user experience, as users may struggle to see the virtual agent's entire body while focusing on the game board (Lee et al., 2018b; Wang et al., 2019). Potential solutions include using miniature virtual agent bodies (Kim et al., 2020; Wang et al., 2019) or scaling up the board to a giant size on the ground, allowing players to walk on it and enhancing visibility and interaction (Bocchi et al., 2024; Mouton et al., 2017).

In our work, we incorporated miniature virtual opponents into a scaled-up Gobang board game (Li et al., 2022) within an MR environment, enabling agents to navigate and place pieces on the enlarged board. Additionally, we introduced the virtual agent as an opponent in the first user study and as an assistant in the second, allowing us to investigate the impact of locomotion and postures on user engagement and social perception in both competitive and cooperative settings.

2.4. EEG for measuring engagement and attention

As reviewed in Section 2.2, EEG is a key physiological measure in human-agent interaction studies, commonly used to assess cognitive load (Antonenko et al., 2010), visual attention (Busch & VanRullen, 2010), and engagement (McMahan et al., 2015), and more. EEG signals encompass brain rhythm frequency bands, including delta (<4 Hz), theta (4–8 Hz), alpha (8–13 Hz) and beta (13–30 Hz), and gamma (>30 Hz) bands (Steven, 2014). Building on these frequency bands, researchers have developed EEG-based indexes to measure user engagement and attention. For example, Berka et al. (Berka et al., 2007) measured engagement by a four-class quadratic discriminant function analysis based on power spectra variables from Fz-POz and Cz-POz. They argued that the EEG engagement index in their study reflects information-gathering, visual processing, and attention allocation. Similarly, Coelli et al. (2015) used beta power divided by alpha power obtained from the frontal and parietal cortex to calculate the engagement index.

The findings validated the effectiveness of the engagement index in reflecting task engagement and further indicated the important role of the frontal cortex in sustained attention and vigilance.

Notably, Pope et al. (1995) developed an EEG-based engagement index calculated as the ratio of combined beta power to the sum of combined alpha and theta power, using signals from the Cz, Pz, P3, and P4 electrodes. They applied it in a closed-loop system to adjust the user task allocation to maintain higher task engagement. Freeman et al. (1999) extended Pope et al.'s work with larger sample sizes and more detailed observations, further validating the effectiveness of this EEG-based engagement index. Since its introduction, this EEG-based engagement index has been widely used to measure human cognitive states in later research (Castiblanco Jimenez et al., 2022; McMahan et al., 2015; Nuamah & Seong, 2018; Rajendran et al., 2022; Szafir & Mutlu, 2012). Furthermore, beyond calculating the EEG-based engagement index using combined power bands, other studies have also utilized single-electrode data from various brain regions, including the frontal, occipital, parietal, and temporal lobes (Alimardani et al., 2021; Kosonogov et al., 2024).

In this paper, we employed the EEG-based engagement index proposed by Pope et al. (1995) to assess participants' cognitive engagement, analyzing the index using both combined power bands and single-electrode data. Unlike physiological measures such as GSR and HRV, which primarily capture autonomic responses and exhibit temporal delays in reflecting arousal, EEG directly measures real-time neural activity associated with cognitive processing. Given that Gobang game is a cognitively demanding but relatively low-stress and slow-paced game, EEG offers a more appropriate means than GSR or HRV for capturing participants' cognitive responses in this context. To our knowledge, this work represents one of the first applications of the EEG engagement index in the study of MiRAs.

In summary, locomotion and body postures are critical non-verbal cues in MiRAs, yet previous research has primarily relied on subjective and behavioral measures to evaluate their effects. Board games, with their structured and rule-based nature, provide an ideal setting for studying human interactions with virtual agents. EEG offers a powerful tool for capturing users' cognitive activities and engagement levels, providing valuable physiological insights. In this work, we combine subjective, behavioral, and physiological (i.e., EEG-based engagement index) measurements to examine how MiRAs' locomotion and postures impact users' social perceptions in a Gobang game environment. Through two user studies, featuring virtual agents as opponents and assistants, we comprehensively investigate the influence of non-verbal cues on user engagement and social perception in both competitive and cooperative scenarios.

3. System overview

Our system consists of an IVA and an MR Gobang game, where the IVA either serves as a competitor or assistant to users in the MR Gobang experience using the HoloLens 2 headset (see Figure 1). Previous research highlights how the limited FoV of the HoloLens 2 can influence user interactions with virtual agents (Lee et al., 2019; Li et al., 2018). It has also been observed that miniature embodied agents walking around tend to be perceived as more approachable and relatable by users (Wang et al., 2019). Additionally, the restricted FoV may prompt increased head movements, potentially introducing more noise into the EEG signals. Therefore, our system features a miniature embodied agent designed to navigate the game board and place virtual pieces in the opponent mode or give suggestions in the assistant mode, simulating a player's actions in a giant game (Bocchi et al., 2024).

In the following subsections, we will describe each study's IVA and MR Gobang game configurations. A more detailed system implementation can be found in Chang et al. (2024).

3.1. The Intelligent virtual agents

In our system, we implemented three different kinds of 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**). These agents were integrated into the MR Gobang game (see subsection 3.2) and illustrated in Figure 1. The IVAs were utilized either as opponents (Figure 1(a)–(c)) or assistants (Figure 1(d)–(f)) within the game.

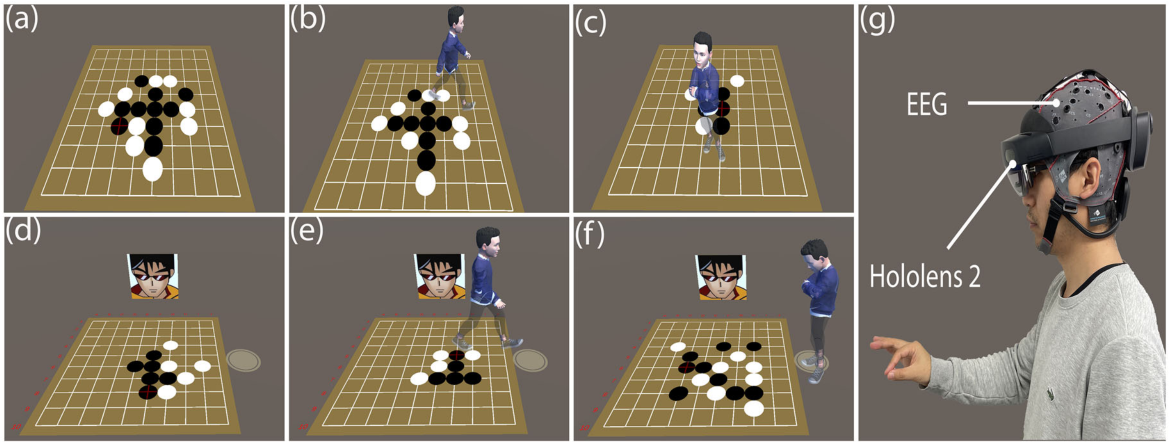


Figure 1. Experimental conditions and hardware setup across both studies: (a) speech (S), (b) speech + locomotion (S + L), and (c) speech + locomotion + posture (S + L + P) from study 1, where the virtual agent acted as an opponent in a turn-based gobang game (adapted from Chang et al. (2024)). (d) S, (e) S + L, and (f) S + L + P from study 2, where the virtual agent served as an assistant, providing strategic hints on demand. (g): hardware setup used in both studies, including the HoloLens 2 for MR interaction and an EEG headset for measuring cognitive engagement (originally presented as (d) in Chang et al. (2024), reordered here for clarity).

IVA **S** (**S**peech): The speech-only agent interacted with users solely through voice with no virtual body (see Figure 1(a) and (d)). In both studies, the virtual agent greeted the players and invited them to play the Gobang game. During gameplay, the virtual opponent verbally urged players to make their move if they delayed for over 6 s, while the virtual assistant offered verbal suggestions for moves when asked. At the end of the game, the virtual opponent expressed happiness if it won, sadness if it lost, and neutrality in the case of a tie. Conversely, the virtual assistant expressed happiness for the player's win, sadness for their loss, and neutrality for a tie. For a detailed explanation of the speech design, see (Chang et al., 2024) for the opponent and Section 6.1 for the assistant.

IVA **S + L** (**S**peech and **L**ocomotion): The speech and locomotion agent was built upon the S condition, with the addition of an embodied virtual agent. This agent featured a fully realized virtual human body and could walk to target points using walking animations. By default, the virtual human faced the user, incorporating subtle body movements to avoid appearing rigid. When walking, the agent first turned its body toward the target location before proceeding. Upon reaching the target, it turned back to face the user. Additionally, the virtual agent could establish and maintain mutual gaze by following the user's gaze, enhancing the realism of the interaction (Gregory et al., 2021). As shown in Figure 1 (b) and (e), the S + L opponent in Study 1 used locomotion to place pieces on the board, while the S + L assistant in Study 2 used locomotion to suggest piece positions.

IVA **S + L + P** (**S**peech, **L**ocomotion and **P**osture): The speech, locomotion, and posture agent was built upon the S + L condition, with the addition of postures to convey emotions and behaviors such as happiness, sadness, and thinking, as well as a neutral expression, as illustrated in Figure 2.

3.2. MR Gobang game

Gobang is a turn-based board game where two players compete, one using white pieces and the other black. Players alternate turns, placing their pieces on the board, aiming to form a line of five consecutive pieces horizontally, vertically, or diagonally to win (see Li et al., 2022 for detailed rules). The possible outcomes for a player are a win, loss, or tie. Gobang strikes a unique balance between simplicity and strategic depth among board games, making it particularly well-suited for studying subtle interactions between users and IVAs without the added complexity of more intricate game mechanics.

We developed an MR Gobang game for HoloLens 2 users, enabling users to play using intuitive hand gestures. Users controlled a virtual cursor on the Gobang board by moving their dominant hand and confirmed their chosen piece position by tapping their thumb and index finger together, causing a virtual piece to appear at the chosen position (see Figure 1(d)). A placement sound accompanied each

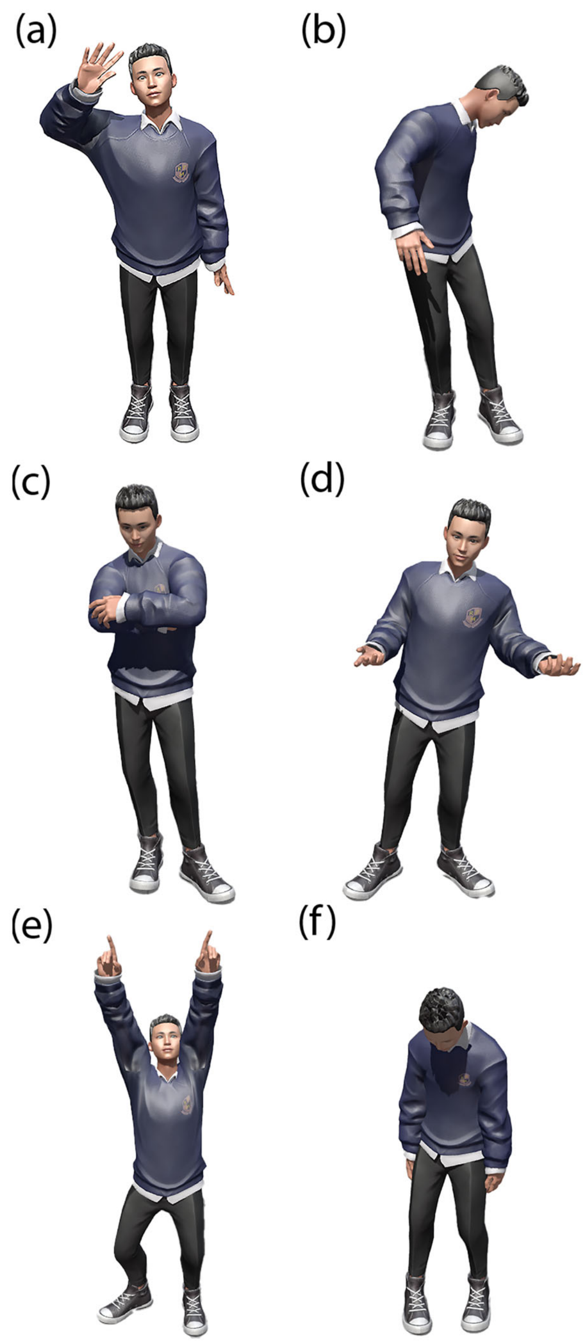


Figure 2. Examples of various postures of the S + L + P agent (Chang et al., 2024): (a) greeting posture at game start, (b) looking around to think about where to place the piece in study 1, (c) folding arms and tapping the left foot to push users in study 1 while watching the gameplay and waiting for a help request in study 2, (d) neutral talking posture when the result is tie in both studies, (e) victory posture to express happy emotion when the agent wins the game in study 1 or when the user win the game in study 2, and (d) looking down to the ground with feet tapping waving back and forth to express sad emotions when the agent loses the game or when the user loses the game in study 2.

piece placement to enhance interactivity. The virtual board, larger than a standard Gobang board, was placed on the ground, resembling a giant board setup. In our studies, participants wearing the HoloLens 2 either competed against the IVA on the virtual board (in Study 1) or received assistance from the IVA (in Study 2).

In Study 1, the virtual opponent's gameplay was driven by Artificial Intelligence (AI) models trained with AlphaZero-Unity[5], incorporating reinforcement learning and Monte Carlo Tree Search (MCTS) (Wang et al., 2021) to generate strategies for competing against human players. Similarly, the same algorithm was used in Study 2 to provide assistant suggestions and manage the computer opponent's moves. For further details on these two studies, refer to Section 5 and Section 6.

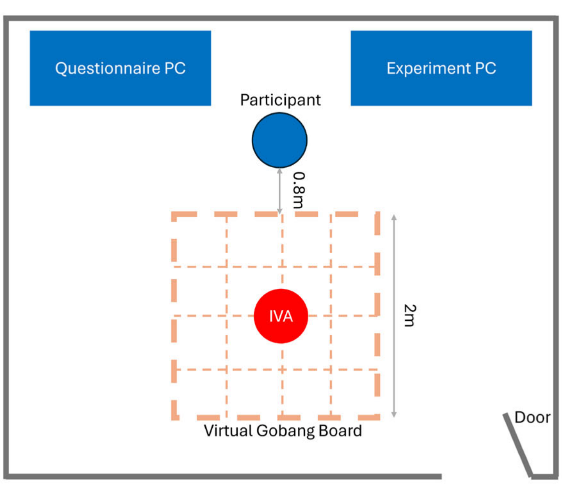
3.3. Implementation

The IVAs and MR Gobang game were developed 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 8 dry EEG electrodes from a Unicorn cap[6] to capture participants' brain signals. The electrodes were repositioned to AF3, P7, P3, O1, O2, P4, Pz, and Cz using a g.tech cap[7] as the Unicorn's original electrode placements did not align with the desired positions, whereas the g.tec cap supported those locations. A more detailed description of system implementation can be found in our previous work (Chang et al., 2024).

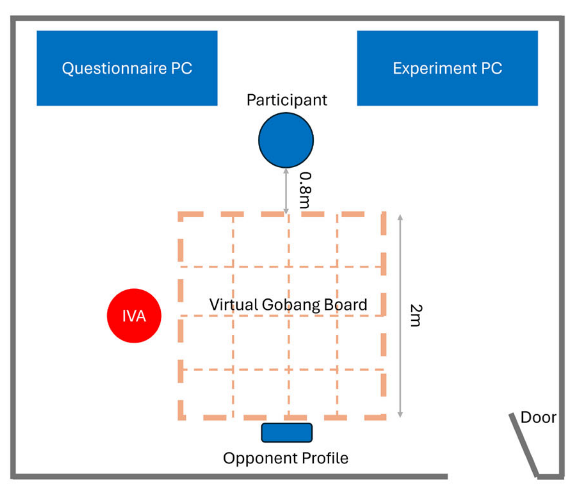
4. Studies overview

Our studies were partially motivated by Leite et al. (2009) and Leite et al. (2010), where the same iCat robot was used in two studies, one as an opponent in a chess game and the other as an empathic companion supporting users during gameplay. Our design was also partially informed by the suitability of the Gobang game as an effective testbed for forming competing and cooperating relationships between users and IVAs. Using the system described in Section 3, we conducted two user studies to examine the impact of locomotion and postures on users' social perceptions within competing and cooperating relationships with IVAs. In these studies, the IVAs served as opponents in Study 1 and assistants in Study 2 in the MR Gobang game.

As shown in Figures 1 and 3, the experimental setup involved a participant equipped with an EEG device and a HoloLens 2 headset, standing approximately 0.8 m away from a virtual 2 m × 2 m square board projected on the floor. In Study 1 (Chang et al., 2024), the embodied opponent began at the center of the virtual board and moved around to place pieces during the game. In Study 2, the embodied assistant initially positioned itself to the right of the board, moved to suggest moves when requested, and returned to its default position afterward. In Study 2, participants played against a computer opponent represented by a profile image of a cartoon character positioned opposite the participants on one side of the chessboard.



(a) Study 1 experimental setup (Chang et al., 2024): The embodied opponent starts at the center of the game board and moves around to place pieces during the game.



(b) Study 2 experimental setup: The embodied assistant defaults to the right of the board, moves to suggest positions when requested, and returns to the default point afterward.

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

Table 1. Measurements used in study 1 and study 2.

Category	Measurements	Study 1: Opponents	Study 2: Assistants
Subjective	Social engagement	Y	Y
	Social presence	Y	Y
	Copresence	Y	Y
	Attentional allocation	Y	Y
	Perceived message understanding	Y	Y
	Perceived affective understanding	Y	Y
	Perceived affective interdependence	Y	Y
	Perceived behavioral interdependence	Y	Y
	Self-Assessment Manikin (SAM) Valence and Arousal	Y	Y
	Rapport	N	Y
	Trust	N	Y
	NASA-TLX Task Load	N	Y
	Interview on Condition Preference	Y	Y
	Interview on Agent Role Preference	Y	Y
	Interview on Emotional Cues Perception	Y	N
Behavioral	Gaze Duration (GD) Proportion	Y	Y
	Task Performance	Y	Y
	Task Completion Time	Y	Y
	Rounds of Play	Y	N
	Player Average Speed	Y	N
Physiological	EEG-based engagement index	Y	Y

Notes: 1. Y indicates the measurement was used in the respective study, while N indicates it was not used.
2. Detailed explanations of the measurements used in Study 1 are in Chang et al. (2024).
3. Those unique to Study 2 are provided in Section 6.

To gain a comprehensive understanding of users' social perceptions of the virtual agents, both studies employed a diverse set of measurements spanning subjective, behavioral, and physiological categories, as outlined in Table 1. The studies primarily focused on the effects of locomotion and postures, investigating their influence on social engagement (Leite et al., 2014), social presence (Harms & Biocca, 2004), and emotional valence and arousal (Lang, 1995). The social presence questionnaire (Harms & Biocca, 2004) we used includes six subdimensions: copresence, attentional allocation, perceived message understanding, perceived affective understanding, perceived affective interdependence, and perceived behavioral interdependence. In Study 2, because the virtual agent was designed as an assistant, we further measured rapport (Cerekovic et al., 2016) and trust (Jian et al., 2000) to evaluate how locomotion and postures influenced rapport and trust building between users and the virtual assistant.

Beyond the questionnaires used in the subjective measurements, semi-structured interviews were conducted in both studies to explore participant preferences for agent designs (i.e., S, S + L, S + L + P) and their preferred agent role (opponent or assistant) and their perceptions of emotional cues from the agents (the latter only in Study 1). To capture role preferences, participants were asked to consider both possible contexts, although they only participated in one study. For example, participants in Study 1 (competitive scenario) were asked whether they would prefer an agent as an opponent, as in the study, or as an assistant providing gameplay suggestions. Similarly, participants in Study 2 (cooperative scenario) were asked whether they would prefer an assistant, as in the study, or an opponent. Behavioral measurements were employed to evaluate the influence of different agent designs on user behavior during the game, while the EEG-based engagement index served as a physiological measure to capture the impact of agent designs on users' brain activity during gameplay.

5. User Study 1: Using MiRAs as opponents

Since Study 1 has been detailed in our previous work (Chang et al., 2024), in this section, we provide a high-level summary of the key results from that work.

To explore the effects of the virtual opponents' locomotion and postures on social presence and engagement, we conducted a within-subjects user study where participants played the MR Gobang game against three different agent conditions (see Figure 1): 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**).

Table 2. Summary of results in study 1 and study 2.

Category	Measurements	Study 1: Opponents	Study 2: Assistants
Subjective	Social engagement	S + L + P > S + L, S	S + L + P > S + L
	SAM Arousal	S + L + P > S	ns
	SAM Valence	ns	S + L + P > S
	Social presence	S + L + P > S + L, S	S + L + P, S + L > S
	copresence	S + L + P > S	S + L + P, S + L > S
	attentional allocation	ns	S + L + P, S + L > S
	perceived message understanding	ns	ns
	perceived affective understanding	S + L + P > S + L	S + L + P > S
	perceived affective interdependence	S + L + P > S	ns
	perceived behavioral interdependence	ns	ns
	Rapport	\	S + L + P, S + L > S
	Trust	\	ns
	NASA-TLX Task Load	ns	\
Behavioral	GD proportion	ns	ns
	Task performance	ns	ns
	Task completion time	ns	ns
	Rounds of play	ns	\
	Player average speed	ns	\
Physiological	EEG-based engagement index	S + L + P < S + L (at AF3)	S + L + P > S (at P3)

Notes: ns indicates no significant difference, \ denotes the measurement was not used, and all other entries indicate significant differences.

Our hypotheses were:

H1.1a: Compared to a speech-only virtual opponent, incorporating locomotion in the virtual agent can enhance human-agent social interaction.

H1.1b: Combining locomotion and postures for the virtual agent can further improve social perception.

H1.2a: Compared to a speech-only virtual opponent, incorporating locomotion in the virtual opponent can increase the EEG-based engagement index.

H1.2b: Combining locomotion and postures for the virtual opponent can further increase the EEG-based engagement index.

We formulated **H1.1**, comprising **H1.1a** and **H1.1b**, based on insights from the literature review (see Section 2.1). Similarly, for **H1.2**, which includes **H1.2a** and **H1.2b**, we hypothesized that if **H1.1** were supported by our experimental design, corresponding significant effects would also manifest in the EEG signals.

As shown in Table 2, the S + L + P opponent demonstrated significantly higher social presence, engagement, and emotional arousal than the S opponent and was the participants' preferred choice. However, the EEG-based engagement index at AF3 was lower in the S + L + P condition compared to the S + L condition, contrasting with the engagement reported in the questionnaires. No significant differences were observed in behavioral measurements. As discussed in our previous work (Chang et al., 2024), we suggest that the locomotion and postures of the S + L + P agent enhance social presence and engagement as reflected in the questionnaires. However, these behaviors may have diverted participants' attention from the game, leading to a lower EEG engagement index, as supported by the behavioral and physiological results. In competitive settings, such distractions could further contribute to reduced task performance.

6. User Study 2: Using MiRAs as assistants

This section focuses on the experimental design, task, measurements, results, and discussions of User Study 2. The experimental setup is detailed in Section 4, and the experimental procedures for Study 2 were identical to those of Study 1, as described in Chang et al. (2024).

6.1. Experimental design

In Study 1, interview results indicated that novice players of the Gobang game preferred the virtual agent in the role of an assistant (Chang et al., 2024). Inspired by this finding, we conducted Study 2, where the virtual agent took on an assistant role specifically for novice players in the Gobang game. In Study 2, we mainly investigated the following research question:

How do the virtual assistant postures and locomotion influence the user's social perception and engagement?

Similar to Study 1 hypotheses, we also made the following hypotheses:

H2.1a: Compared to a speech-only virtual assistant, incorporating locomotion in the embodied virtual assistant can enhance human-agent social interaction.

H2.1b: Combining postures and locomotion for the embodied virtual assistant can further improve social perception.

H2.2a: Compared to a speech-only virtual assistant, incorporating locomotion in the embodied virtual assistant can increase the EEG-based engagement index.

H2.2b: Combining postures and locomotion for the embodied virtual assistant can further increase the EEG-based engagement index.

We conducted a within-subjects user study with 20 participants (10 male, 10 female) aged 18 to 33 years ($M = 25.5$, $SD = 4.31$) from the same university campus as Study 1 through campus posters and personal invitations, yet none of the participants in Study 2 had participated in Study 1. We applied the same recruitment criteria as in Study 1 with an extra requirement for participants to be novices in the Gobang game. Regarding prior XR experiences, nine participants reported no experience using hand interactions with virtual content in AR/VR/MR environments; eight reported little experience (i.e., once every few months or annually), and three reported frequent use (i.e., daily, weekly, or monthly). Concerning IVAs like Amazon Alexa, Microsoft Cortana, and Apple Siri, ten participants reported little experience, seven reported frequent use, and three had never used such agents. We employed a balanced Latin Square design to counterbalance the order of conditions and reduce learning effects. The entire experiment lasted approximately 90 min, and participants were thanked for their time and contribution with a $20 gift card in local currency.

We followed the same procedure as Study 1 (Chang et al., 2024), with the addition of a Wizard-of-Oz method (Dahlbäck et al., 1993) to detect participants' requests for help and activate the virtual assistant's responses. Before the study, participants were instructed to request assistance by speaking any sentence that included the virtual assistant's name, "Alex." During the study, any time when participants asked for help, the experimenter pressed the space key slightly on the keyboard to trigger the virtual assistant's helping response. To ensure the virtual assistants' suggesting behaviors were triggered, all participants were required to request assistance at least once during the session. Similar to the Study 1 design, we also designed three similar agent conditions in Study 2:

S: **S**peech-only condition, where the virtual assistant responded to participants only through speech. At the beginning of the game, the virtual assistant greeted the participant and introduced himself as an assistant. During the game, when the participant asked for help, the virtual assistant provided a hint for the next move. For example, when the participant asked for help by saying, "*Alex, where should I place my next piece?*" the virtual assistant would hint at the next move by saying, "*Based on the current situation, I suggest you place a piece at B5.*" The virtual assistant used the same AI algorithm as in Study 1 to calculate the suggested next move. At the end of the game, the assistant would express empathic emotions verbally based on the game results. For example, if the participant won, the virtual assistant would express happiness by saying "*Yes, we won! I am so happy about our victory.*" If the participant lost, the virtual assistant would express sadness and encourage the participant by saying "*I'm a bit disappointed about our loss, but don't worry, I know we can improve and succeed next time!*" If drawn, the virtual assistant would express satisfaction by saying "*It's a draw! I'm satisfied with our performance.*" As in the opponent agent design in Study 1, all speech includes multiple synonymous sentences to prevent stiffness.

S + L: **S**peech and **L**ocomotion condition, where the virtual assistant could move toward the suggested grid before providing verbal suggestions. At the beginning of the game, the virtual assistant faces and looks at the participant during the greeting and introduction. During gameplay, the assistant focused on the board, turning its body toward newly placed pieces as if observing the game. If participants looked at the assistant while it was in its initial position, the assistant would look back at them. When the participant requested help, the assistant moved toward the suggested grid, focused on it, and then provided verbal suggestions. After offering the suggestion, the assistant returned to its original position.

S + L+P: **S**peech, **L**ocomotion, and **P**osture condition, where the virtual assistant could further express postures based on the S + L condition. As shown in Figure 2(a), the virtual assistant could wave hands to greet users and make a welcoming gesture at the greeting and introduction phase. While the virtual assistant is watching the gameplay, it could randomly switch its postures from idle animation with slight body movement to a waiting posture with arms folded and head tilted (see Figure 2(c)). While providing suggestions, the virtual assistant could also exhibit a talking gesture by moving its head and arms slightly, as demonstrated in Figure 2(d). At the end of the game, the virtual assistant displayed emotions based on the outcome: happiness for the user's win, sadness for their loss, and neutrality for a tie, using emotional postures as shown in Figure 2.

6.2. Experimental task

As shown in Figure 1(d)–(f), participants played the MR Gobang game with the assistance of different virtual assistants under different study conditions. When the game started, the virtual assistant initiated the interaction by greeting participants and inviting them to the game. Once the participants verbally agreed to play the game, the computer opponent always made the first move with black pieces. During the game, participants played the MR Gobang game against a computer opponent, represented by a cartoon character profile image positioned opposite them. when participants asked for assistance, the virtual assistant verbally suggested positions to participants with different behaviors expressed based on the experimental conditions. At the end of the game, the virtual assistant verbally expressed empathy with the participants' outcomes, following the different designs outlined in Section 6.1.

6.3. Measurements

The measurements in Study 2 were similar to those in Study 1, with some modifications. In the following subsections, we detail the differences in measurements between two studies.

6.3.1. Subjective measurements

For the subjective measurements, we used the same social engagement and SAM valence-arousal questionnaires as in Study 1. Because the social presence questionnaire used in Study 1 contains 36 questions (Harms & Biocca, 2004), we opted for a shorter version in Study 2. Specifically, we adopted a simplified social presence measure with two items per subscale from (Leite et al., 2009) to ensure the overall questionnaire remained concise. We removed the NASA-TLX questionnaire because no significant difference was found in Study 1. Instead, we further added rapport (Cerekovic et al., 2016) and trust (Jian et al., 2000) questionnaires to understand how the virtual assistant's behavior influenced the user's rapport and trust. During the interview, we also explored participants' preferences for different agent types (i.e., S, S + L, or S + L + P) and roles (i.e., assistant or opponent).

6.3.2. Physiological measurement

We used the same EEG-engagement index measurement as Study 1 (Chang et al., 2024) to explore how the agent postures and locomotion influenced users' brain activity when the agent served as an on-demand assistant. We applied a bandpass filter ($1 \sim 40$ Hz) to reduce the noise and artifacts in the EEG signals, followed by manually check each participant's EEG signals per condition and removing bad data segments using the MNE-Python library (Gramfort et al., 2013). In total, 3.2% of the data was identified as bad and subsequently removed. We applied baseline correction[8] to nomalize the EEG data using the 1-second data before the agent starts greeting the users as a baseline. Later, we extracted the data related to the game task between agent greeting user and game end based on the EEG markers set during the runtime. We further calculated the Power Spectral Density (PSD) for each frequency band in the EEG-engagement index equation 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. More details about EEG data analysis and EEG markers can be found in our previous work (Chang et al., 2024).

6.3.3. Behavioral measurements

Regarding behavioral measurements, we retained only GD proportion, task performance, and task completion time, excluding the player's average speed and number of rounds played. This decision was based on the following reasons: First since the assistant was available on-demand rather than as a turn-based opponent, the player's average speed was less relevant. Second, both the number of rounds played and task completion time reflect the overall duration of interaction, so we opted to keep only the task completion time. Furthermore, instead of logging the number of rounds played, we recorded the number of assistance requests from players to examine how different agent designs influence users' willingness to seek assistance.

6.4. Results of Study 2

6.4.1. Subjective measurements results

Social engagement. The Shapiro-Wilk test showed the engagement questionnaire data met the normality assumption. However, Mauchly's sphericity test showed the assumption of sphericity was violated. As shown in Figure 4, the Friedman test indicated there was a significant difference among the three agent conditions (χ^2 (2) = 8.70, $p = 0.013$) with small effect size ($W = 0.218$). 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 + L condition (Z =–1.977, $p = 0.048$). No other significant difference was found in other pairwise comparisons.

Social presence. The Shapiro-Wilk test showed that the social presence data in S ($p = 0.853$), S + L ($p = 0.737$) and S + L + P ($p = 0.330$) were all normally distributed. Mauchly's test indicated the sphericity assumption was also met ($W = 0.732$, $p = 0.06$).

As shown in Figure 5, we saw overall a significant difference in average Social Presence scores across the three experimental conditions using the repeated ANOVA ($F(2,38) = 12.54$, $p < 0.0001$, $\eta_g^2 = 0.1$). The Post-Hoc pairwise t-tests with Bonferroni correction showed that the feeling of Social Presence in the S condition was significantly lower than that in the S + L condition ($t(19) = -5.26$, $p < 0.001$) and S + L + P condition ($t(19) = -3.99$, $p = 0.002$). No significant difference was found between the S + L + P and S + L conditions.

Copresence. The Shapiro-Wilk test showed the copresence data in S + L ($p < 0.01$) and S + L + P ($p = 0.578$) violated the normality assumption. As shown in Figure 6(a), the Friedman test showed there was a significant difference among the three agent conditions (χ^2 (2) = 21.34, $p < 0.0001$), with a large effect size ($W = 0.534$) by the following Kendall's W test. The post hoc Wilcoxon signed-rank test with Bonferroni correction revealed the copresence in the S condition was significantly lower than that in the S + L condition ($Z = -3.291$, $p = 0.001$) and S + L + P condition ($Z = -2.968$, $p = 0.003$). No significant difference was found between S + L and S + L + P.

Attentional allocation. The Shapiro-Wilk test showed the attentional allocation data in S ($p = 0.049$) and S + L + P ($p = 0.003$) conditions violated the normality assumption. As shown in Figure 6(b), the Friedman test showed there was a significant difference between the three agent conditions (χ^2 (2) = 19.2, $p < 0.0001$), with a moderate effect size ($W = 0.48$) by the following Kendall's W test. The post hoc Wilcoxon signed-rank test with Bonferroni correction revealed that the copresence in the S condition was significantly lower than that in the S + L condition ($Z = -2.878$, $p = 0.004$) and S + L + P condition ($Z = -2.748$, $p = 0.006$). No significant difference was found between S + L and S + L + P.

Perceived message understanding. The Shapiro-Wilk test showed the perceived message understanding data in S + L ($p = 0.014$) and S + L + P ($p = 0.049$) conditions violated the normality assumption. As shown in Figure 6(c), the Friedman test showed no significant difference among the three agent conditions.

Perceived affective understanding. The Shapiro-Wilk test showed the perceived affective understanding data violated the normality assumption in S condition ($p = 0.011$). As shown in Figure 6(d), the Friedman test showed there was a significant difference between the three agent conditions

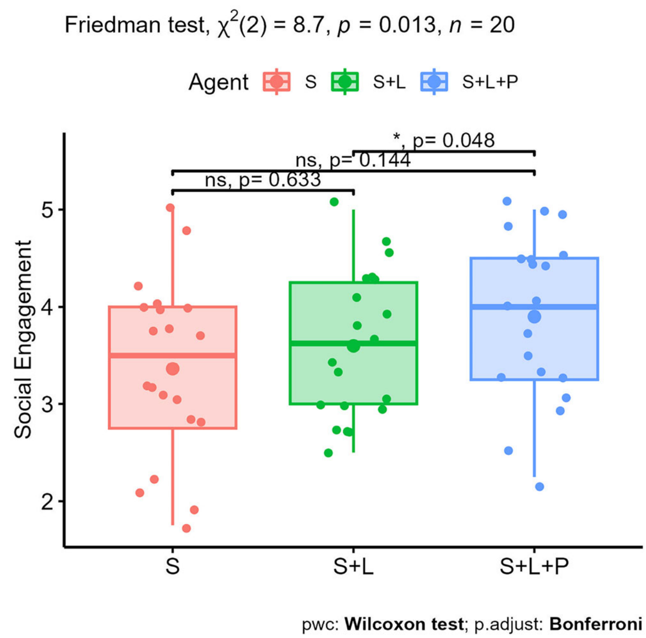


Figure 4. Results of the engagement questionnaire. (5-point likert Scale from 1 to 5; the larger solid dot in each condition represents mean value; Statistical significance: *($p < 0.05$)).

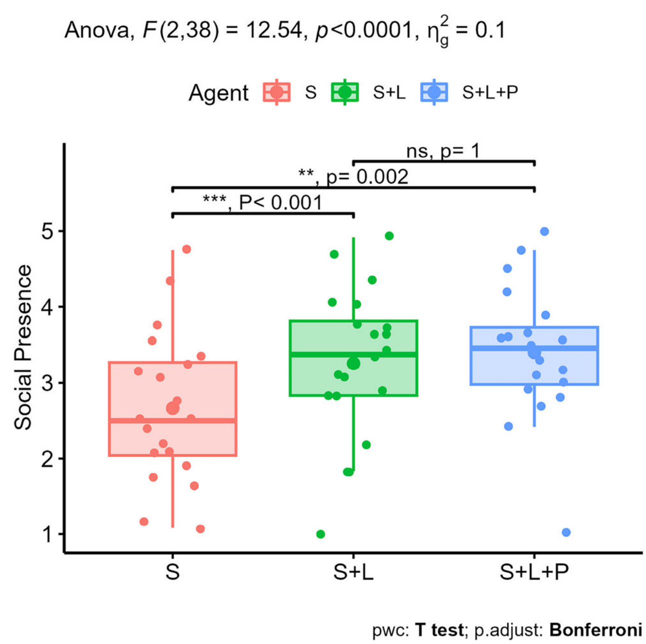
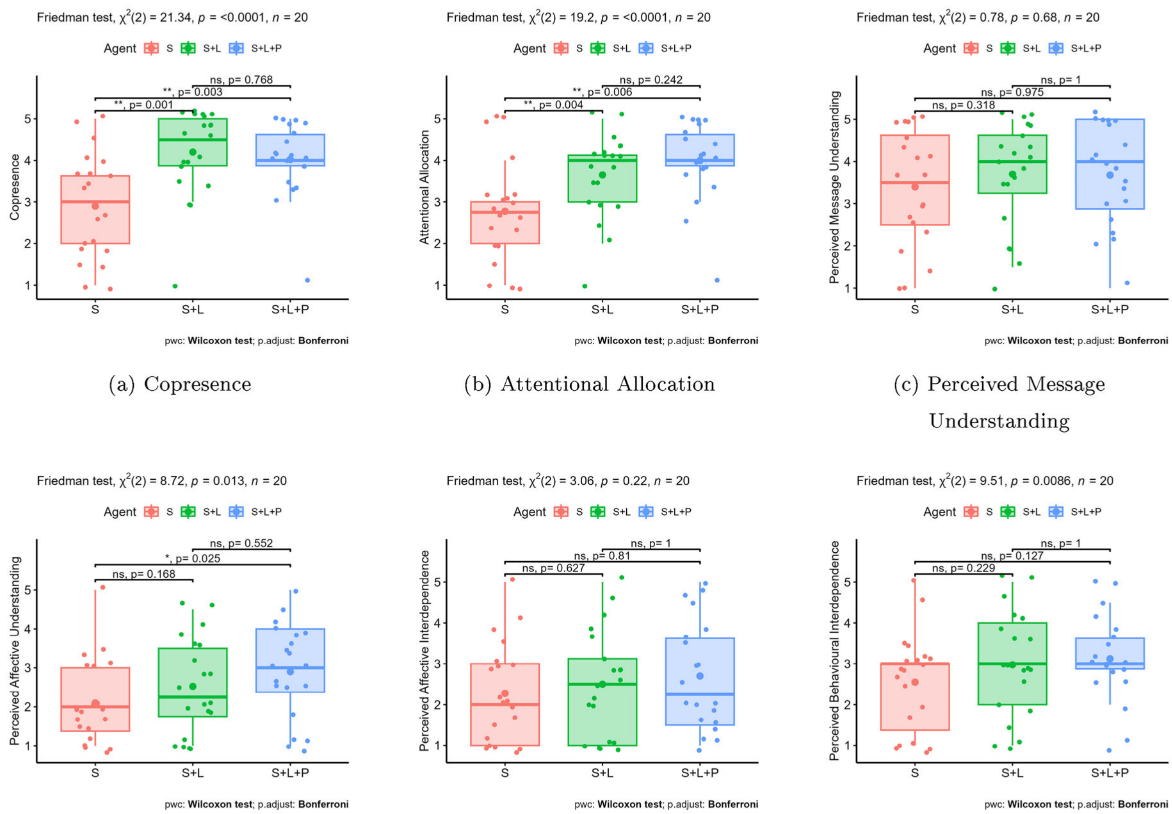


Figure 5. Results of the social presence. (5-point likert Scale from 1 to 5; the larger solid dot in each condition represents mean value; Statistical significance: * ($p < 0.05$), ** ($p < 0.001$)).

(χ^2 (2) = 8.72, $p = 0.013$), with a small effect size ($W = 0.218$) by the following Kendall's W test. The post hoc Wilcoxon signed-rank test with Bonferroni correction revealed the copresence in the S + L + P condition was significantly higher than that in the S condition ($Z = -2.241$, $p = 0.025$). No other significant difference was found.

Perceived affective interdependence. The Shapiro-Wilk test showed the perceived affective Interdependence data violated the normality assumption in S condition ($p = 0.026$). As shown in Figure 6(e), no significant difference was found among the three conditions using the Friedman test.

Perceived behavioural interdependence. The Shapiro-Wilk test showed that the perceived affective Interdependence data violated the normally assumption in S condition ($p = 0.039$). As shown in



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

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

Figure 6. 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: * ($p < 0.05$), ** ($p < 0.001$).

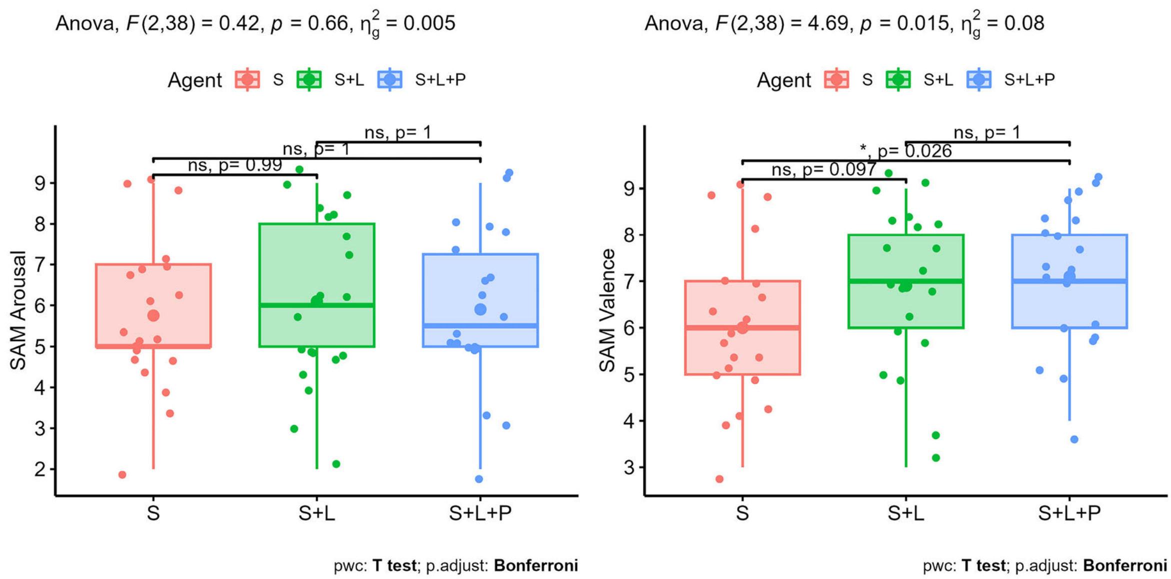
Figure 6(f), although the Friedman test showed an overall significant difference among the three conditions (χ^2 (2) = 9.51, $p = 0.009$), the post hoc Wilcoxon signed-rank test with Bonferroni correction revealed no significant difference was found in the pairwise comparisons.

SAM valence and arousal. For the SAM Arousal scale, the Shapiro-Wilk test showed the normality assumption was met in all agent conditions (S ($p = 0.179$), S + L ($p = 0.204$), S + L + P($p = 0.224$)). As shown in Figure 7(a), no significant difference was found across the three agent conditions.

For the SAM Valence scale, the Shapiro-Wilk test showed the normality assumption was met in all agent conditions (S ($p = 0.175$), S + L ($p = 0.081$), S + L + P($p = 0.131$)). Mauchly's test ($W = 0.998$, $p = 0.984$) did not indicate any violation of sphericity. As shown in Figure 7(b), we saw overall a significant difference in average SAM valence scores across the three experimental conditions using the repeated ANOVA (F(2,38) = 4.69, $p = 0.015$, $\eta_g^2 = 0.08$). The Post-Hoc pairwise t-tests with Bonferroni correction showed that the SAM valence in the S + L + P condition was significantly higher than that in the S condition (t(19) = −2.92, $p < 0.026$). No other significant difference was found.

Rapport. The Shapiro-Wilk test showed the rapport data in S ($p = 0.227$), S + L ($p = 0.478$), S + L + P($p = 0.097$)) met the normality assumption. Mauchly's test ($W = 0.88$, $p = 0.317$) did not indicate any violation of sphericity. As shown in Figure 8(b), a repeated ANOVA (F(2,38) = 6.19, $p = 0.005$, $\eta_g^2 = 0.1$) revealed there was significant difference in average rapport scores between the three agent conditions. The Post-Hoc pairwise t-tests with Bonferroni correction showed that the rapport in the S condition was significantly lower than that in the S + L condition (t(19) = −3.29, $p < 0.011$) and S + L + P (t(19) = −2.79, $p < 0.035$). No significant difference was found between S + L and S + L + P condition.

Trust. The Shapiro-Wilk test showed the trust data in S ($p = 0.481$), S + L ($p = 0.844$), S + L + P($p = 0.765$)) met the normality assumption. However, Mauchly's test ($W = 0.644$, $p = 0.046$)



(a) SAM Arousal (b) SAM Valence

Figure 7. 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: * ($p < 0.05$).

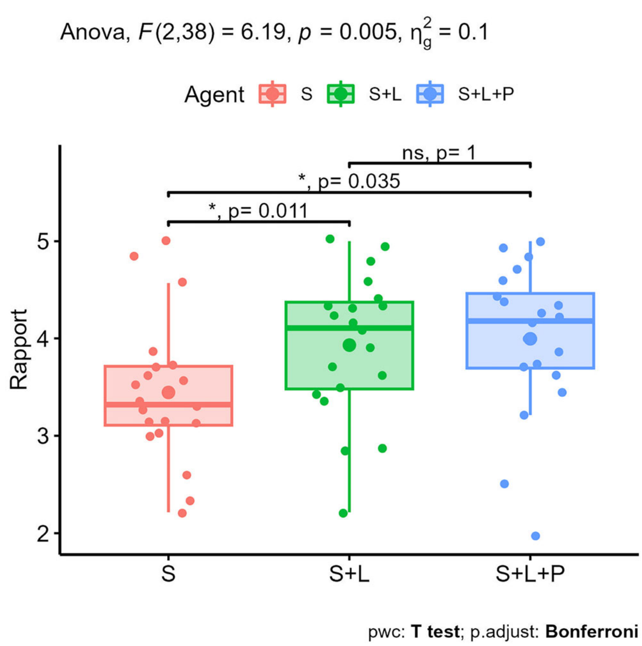


Figure 8. Results of the rapport questionnaire. (5-point likert Scale from 1 to 5; the larger solid dot in each condition represents mean value; Statistical significance: * ($p < 0.05$)).

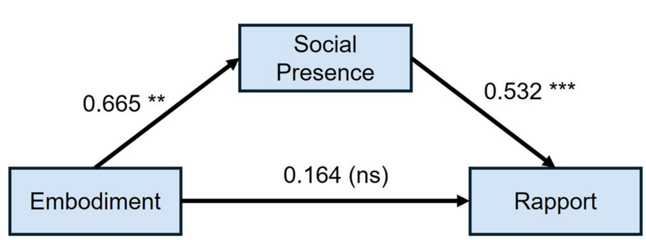


Figure 9. Mediation analysis diagram showing the relationship between embodiment and rapport, mediated by social presence. (Statistical significance: ** ($p < 0.01$), *** ($p < 0.001$)).

indicated the sphericity assumption was violated. A repeated ANOVA ($F(2,38) = 0.98$, $p = 0.38$, $\eta_g^2 = 0.02$) revealed there was no significant difference across the three agent conditions on trust. Descriptive statistics indicated that all three agent conditions had trust scores near the neutral value of 5. However, the S + L + P condition ($M = 5.42$, $SD = 1.34$) and the S + L condition ($M = 5.51$, $SD = 1.05$) showed slightly higher trust scores compared to the S condition ($M = 5.16$, $SD = 1.16$).

Mediation analysis. Given that the S + L + P and S + L conditions demonstrated significantly higher social presence (see Section 6.4.1) and rapport (see Section 6.4.1) compared to the S condition, we conducted a mediation analysis to test whether social presence mediates the relationship between embodiment and rapport.

For this purpose, we combined the S + L + P and S + L conditions into a single embodiment group (totaling 40 participants) and compared it to the S group (speech-only condition) consisting of 20 participants. We acknowledge that the unequal sample sizes between the embodiment and S groups could potentially influence the results by affecting the statistical power and the stability of the estimates. To mitigate this concern, we applied bootstrapping (Aguinis et al., 2017) with 5000 resamples in the mediation analysis using PROCESS (Model 4) (Hayes, 2022).

As shown in Figure 9, the results indicated that embodiment significantly predicted social presence ($b = 0.665$, $p = 0.013$), and social presence significantly predicted Rapport ($b = 0.532$, $p < 0.001$). The direct effect of embodiment on rapport was not significant ($b = 0.164$, $p = 0.307$), but the indirect effect through social presence was significant ($b = 0.354$, 95% CI = [0.0741, 0.6653]). These results indicate that the effect of embodiment on rapport is fully mediated by social presence.

Interview on user preference. Consistent with the procedure in Study 1, participants ranked their most and least preferred conditions after interacting with all three agent conditions. A Chi-Square Goodness of Fit test yielded a significant difference for the most liked agent type (χ^2 (2) = 24.258, $p < 0.0001$) as well as the least liked agent type (χ^2 (2) = 13.65, $p = 0.001$) (see Figure 10(a)).

In addition to identifying their most and least preferred agent types, participants were asked whether they preferred playing with an assistant or an opponent to improve their Gobang game skills. As shown in Figure 10(b), participants in Study 2 exhibited a stronger preference for playing with an assistant over an opponent (χ^2 (1) = 4.950, $p = 0.026$), aligning with the findings from Study 1 (Chang et al., 2024).

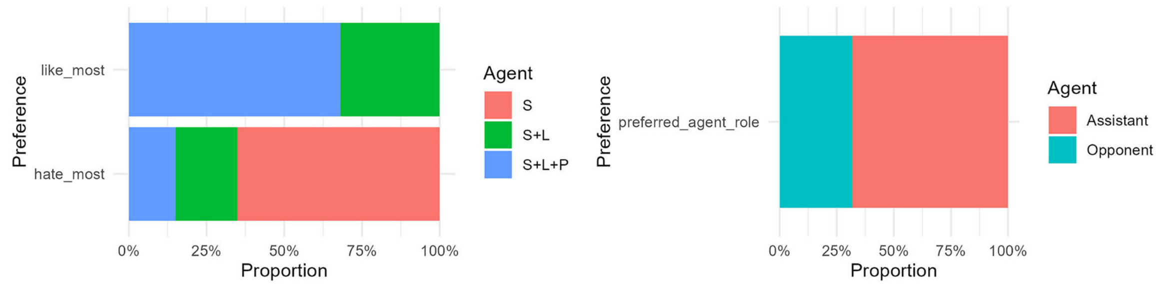
6.4.2. Physiological measurement results

EEG-based engagement index. Due to technical issues, the EEG data were incorrectly recorded for four participants P03, P05, P09, and P13. As a result, data from two female and two male participants were omitted while processing the EEG results. The Shapiro-Wilk test showed the EEG-based engagement index at P3 in S ($p = 0.544$), S + L ($p = 0.126$), S + L + P($p = 0.095$)) met the normality assumption. Mauchly's test ($W = 0.742$, $p = 0.068$) did not indicate any violation of sphericity. As shown in Figure 11, the Friedman test showed there was a significant difference in engagement between the three agent conditions (χ^2 (2) = 9.88, $p < 0.007$), with a moderate effect size ($W = 0.309$) by the following Kendall's W test. The post hoc Wilcoxon signed-rank test with Bonferroni correction revealed the copresence in the S + L + P condition was significantly higher than that in the S condition ($Z = -2.878$, $p = 0.004$). Note that there was a trend toward significance ($p = 0.064$) between the S and S + L conditions, where S + L showed higher values than S, suggesting a potential effect that may be confirmed with a larger sample size. No other significant difference was found regarding the EEG-based engagement index.

As in Study 1, we also checked the correlation between the EEG-based engagement index and social engagement. However, the Pearson correlation test showed no significant correlation between the social engagement and EEG-based engagement index in the S ($r = -0.375$, $p = 0.151$), S + L ($r = -0.06$, $p = 0.826$) and S + L + P ($r = -0.005$, $p = 0.985$) conditions.

6.4.3. Behavioral measurement results

Number of assistance requests. The Shapiro-Wilk test indicated the number of assistance requests in the S + L + P condition ($p = 0.02$) broke the normality assumption. No significant difference was found across the three agent conditions. Descriptive statistics showed the number of assistance requests was



(a) Agent behavior preference (participants showed significant preferences among the three agent conditions for the 'like most' question, but no significant differences for the 'hate most' question.)

(b) Agent role preference (Participants with higher Gobang skill levels preferred the opponent role, while participants with lower skill levels preferred the assistant role.)

Figure 10. Agent behavior and role preference.

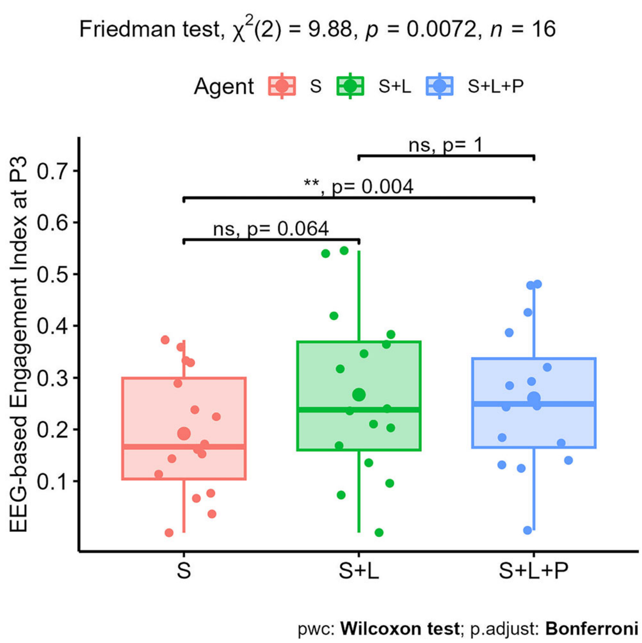


Figure 11. Results of the EEG-based engagement index at P3 (the larger solid dot in each condition represents mean value; Statistical significance: ** ($p < 0.01$).

close in all three agent conditions with S ($M = 4.90\%$, $SD = 3.60\%$), S + L ($M = 4.10\%$, $SD = 3.29\%$) and S + L + P ($M = 4.95\%$, $SD = 3.17\%$)

GD proportion. Due to technical issues with the eye-tracking equipment, we failed to record eye gaze data for four participants (P02, P03, P04, and P13). Consequently, eye gaze data from three female and one male participant were omitted. Therefore, we only had 16 participants' eye gaze data. The Shapiro-Wilk test indicated that eye gaze duration proportion in both the S + L ($p = 0.993$) and S + L + P ($p = 0.204$) conditions were normally distributed. No significant difference was found in average eye gaze duration between these two groups using a pairwise t-test with Bonferroni correction ($t(15) = -0.287$, $p = 0.778$). Descriptive statistics showed the S + L + P ($M = 16.371\%$, $SD = 6.946\%$) agent received similar eye gaze duration proportion to the S + L ($M = 15.732\%$, $SD = 6.332\%$) agent.

Task performance. 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 ($\chi^2(4) = 4.739, p = 0.315$). 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

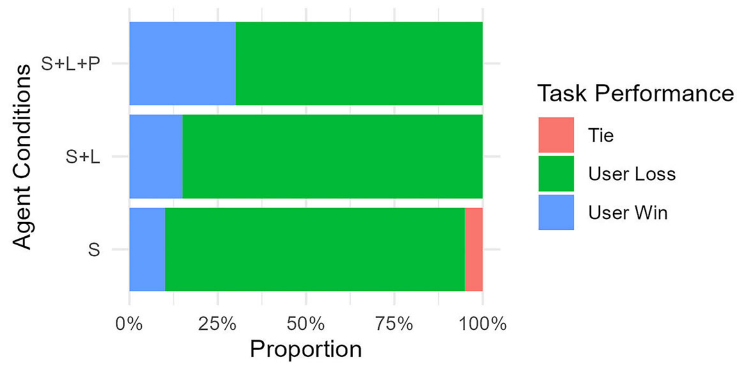


Figure 12. Results of task performance.

the S condition ($\chi^2(2) = 24.1, p < 0.001$) and the S + L condition ($\chi^2(1) = 9.8, p = 0.002$), while no significant difference was observed in the S + L + P condition ($\chi^2(1) = 3.2, p = 0.07$).

Task completion time. The Shapiro-Wilk test showed that the normality assumption of task completion time was violated in all three agent conditions. No significant difference was found in the task completion time between the three conditions using a Friedman test ($\chi^2(2) = 0.4, p = 0.819$). Descriptive statistics showed the S + L + P ($M = 264.50$, $SD = 162.18$) and S + L ($M = 211.55$, $SD = 114.53$) conditions had shorter task completion time on average than the S ($M = 293.20$, $SD = 223.07$) condition.

6.5. Discussion of Study 2

Consistent with Study 1, we observed significant differences across the three agent conditions in subjective and physiological measurements, but no significant differences in behavioral measurements. The social presence and rapport results directly supported **H2.1a**: both the S + L + P and S + L agent conditions were rated significantly higher than the S condition. Only the social engagement results directly supported **H2.1b**, with the S + L + P agent receiving higher ratings than the S + L agent. Additionally, the S + L + P agent was rated significantly higher than the S agent condition in SAM valence, rapport, and social presence, partially supporting **H2.1b**. The EEG-based engagement index partially supported **H2.2a** but not **H2.2b**, as we only found that the EEG-based engagement index at P3 was significantly higher in the S + L + P agent condition than the S condition, while no other significance was found in EEG-based engagement index.

6.5.1. Postures and locomotion in virtual assistants build rapport, but not trust

Our hypothesis **H2.1a** was supported by the social presence questionnaire results. The embodied agents (S + L + P and S + L conditions) received significantly higher average social presence scores than the speech-only agent (S condition). Notably, the same trend was also observed in social presence's copresence and attentional allocation subscales. These results aligned with the findings presented in Kim et al. (2018), who reported that virtual assistants embodied in MR environments incorporating gestures, or both gestures and locomotion, enhanced social and spatial presence.

Our results of the rapport questionnaire also supported **H2.1a**. Both S + L + P and S + L agent conditions had significantly higher rapport than the S agent conditions. As highlighted in Huang et al. (2011), virtual humans' appropriate verbal and non-verbal behaviors could enhance rapport during interaction with human users. Our findings on rapport contribute to understanding virtual rapport-building, particularly in MR environments, where incorporating postures and locomotion in virtual assistants can facilitate rapport.

Moreover, our mediation analysis revealed that the effect of embodiment on rapport was fully mediated by social presence. Specifically, the embodied agent captured more user attention and enhanced the sense of copresence, key factors reflected in the copresence and attentional allocation subscales of social presence. This, in turn, led to higher rapport during the assistance task. These findings clarify the mechanism of virtual rapport-building, underscoring the pivotal role of social presence in fostering rapport with virtual assistants in MR.

In addition to social presence and rapport, the results of the SAM valence measure partially supported **H2.1a**. Specifically, the S + L + P condition showed significantly higher average valence values than the S condition, while no other significant differences were observed. This result aligns with the perceived affective understanding subscale of social presence, indicating that users experienced stronger positive emotional cues from the virtual agent in the S + L + P condition. For example, P06 mentioned, "*… the agent with postures was more active so I could capture more emotional cues. It could show thinking behaviors while watching the game and express emotional postures at the game end …* " and P07 said, "*… the one with gestures was good because it felt easy to communicate with and more approachable …* " Moreover, the S + L + P condition was further shown to be significantly higher than the S + L in average social engagement scores, which directly supported **H2.1b**. This could be because the postures in the S + L + P were effective in maintaining users' attention (Morel & Ach, 2011) and conveying emotional cues (Randhavane et al., 2019b). For instance, P18 commented, "*I like the agent with postures because it makes me focus more on the interaction during the game … I paid more attention to him … I looked at his face while he was talking …* " P20 said, "*I like the second one most (i.e., S + L + P for this participant). It felt more like a human … it's easier to interact with him and easier to understand what he is trying to say as well.*"

Despite the significant differences discussed above, we couldn't find any significant difference in the trust questionnaire results. The descriptive statistics of the average trust scores showed all three agent conditions had a trust score close to a neutral value of 5. This may be because participants were novices in the Gobang game and, therefore, less capable of accurately evaluating the virtual assistants' suggestions. For example, P12 said, "*… I was never skeptical of the agent …* "

A systematic review by Rheu et al. (2021) on trust-building factors in conversational agents also noted that embodiment can increase trust, but not consistently. They found that virtual agent movement is not always critical for enhancing trust and emphasized that the agent's ability is a key factor in fostering trust. In Study 2, the virtual assistants' suggestions were generated using the same algorithm as the Study 1 opponent, proven effective against novice players. Consequently, the virtual assistants consistently provided reasonable suggestions across all conditions, maintaining a consistent level of trust across assistant types.

In summary, our findings show that postures and locomotion in virtual assistants significantly enhance social presence, rapport, and engagement in MR environments, supporting **H2.1a** and **H2.1b**. Embodied agents (S + L + P and S + L) captured user attention and conveyed emotional cues more effectively than the speech-only agent (S). However, trust remained neutral across all conditions, likely due to the participants' novice status and the assistance system's strong capability to provide reasonable suggestions. These results highlight that while combining postures and locomotion fosters rapport and engagement, trust depends more on the assistant's ability than its virtual body movement.

6.5.2. The EEG-based engagement index complements attention measurements for the virtual assistant in the board game

Our EEG-based engagement index results partially supported **H2.2a** but did not support **H2.2b**. Specifically, we found that the EEG-based engagement index at electrode P3 in the S + L + P condition was significantly higher than in the S condition. P3, located in the left parietal lobe, is associated with attentional processing (Behrmann et al., 2004; Chang et al., 2022) and motor attention (Rushworth et al., 2001, 1997). In earlier work, Ganesh et al. (2012) demonstrated that the left parietal lobe is involved in self-identification based on agent movement observed from a third-person perspective. Therefore, we attribute the increased engagement in the S + L + P condition to the agent's motor movements, which captured more attention and were reflected in the left parietal lobe activity. The results of the attentional allocation and copresence subscales of social presence could also support this.

Although we did not find a significant difference, the S + L condition showed a higher EEG-based engagement index at P3 than the S condition, reaching marginal significance ($p = 0.064$) with a moderate effect size. This trend toward significance suggests a potential effect that may be confirmed with a larger sample size. Considering this assumption, together with the significant difference between the S + L + P and S conditions in the EEG-based engagement index at P3, social presence, and rapport, we suggest that higher social perception of the agent may contribute to a higher EEG-based engagement index in this Gobang game context. This could be because the embodied agent could capture more users' attention while not causing distraction. We will elaborate on this by integrating behavioral measurements, EEG data, and questionnaire results.

Consistent with the results of the attentional allocation subscale of social presence, the GD proportion did not show a significant difference between the S + L and S + L + P conditions. This suggests that users' visual attention to the virtual assistant's body was similar in these two conditions. This similarity may be due to the comparable number of assistance requests across all three agent conditions. For example, P10 said, *"I felt the two embodied agents were better … I didn't notice the differences. I just felt they were interesting and friendly because they were standing there watching the game … "* When users requested assistance, they either actively looked at the virtual assistant or the embodied agents in the S + L and S + L + P conditions moved into their field of view, prompting passive attention.

However, the GD proportion and the attentional allocation subscale of the social presence questionnaire have limitations in capturing users' overall attention. The GD proportion is derived from HoloLens 2 eye-tracking data, which simplifies gaze direction by projecting two laser lines and determining the gaze target with a single dot (Plopski et al., 2022). The virtual assistant was positioned to the side of the game board, so users focusing on the board might not have their gaze target dot directly on the assistant. Nevertheless, the assistant could still be within users' peripheral vision, and its body and posture might still be seen without direct gaze. The attentional allocation subscale in Study 2 consists of only two items and measures users' general impression of how much they focused attention on the agent or how much the agent focused attention on them. Since users completed this subscale after each session, it reflects a retrospective summary rather than their real-time attention state. In summary, the GD proportion captures only direct gaze, overlooking peripheral attention, while the attentional allocation subscale reflects general impressions of mutual attention rather than real-time attention during the interaction. Therefore, neither measure fully captures users' overall attention during the game.

As discussed at the beginning of this section, the EEG-based engagement index at P3 reflects users' attentional processing during their interaction with the virtual agent. The S + L + P condition showed a significantly higher EEG-based engagement index than the S condition, while the S + L condition showed a marginally higher EEG-based engagement index than the S condition. Similarly, both the S + L + P and S + L conditions had significantly higher attentional allocation scores in the social presence questionnaire. In contrast, the GD proportion did not significantly differ between the S + L + P and S + L conditions. Combining these findings with the discussion in Section 6.5.1, the embodied agents (S + L + P and S + L) attracted more user attention, enhancing social presence and building rapport.

In conclusion, in Study 2, the EEG-based engagement index at P3 measured users' overall attentional allocation during the game. In comparison, the GD proportion measured direct gaze, overlooking peripheral attention, while the attentional allocation subscale captured general impressions of mutual attention rather than moment-to-moment attention. Therefore, the EEG-based engagement index complements attention measurements for the virtual assistant in the board game, providing a more comprehensive understanding of user attention during interactions.

7. Summary of findings, design implications, and limitations

This work investigates the impact of virtual agents' locomotion and postures on social perceptions in MR through a Gobang board game. Study 1 investigated the impact of locomotion and postures in virtual agents acting as opponents, while Study 2 explored their impact with virtual agents serving as assistants. We first synthesize the key findings from both studies and highlight their divergences before discussing design implications and limitations.

7.1. Summary of findings

In both studies, we found that the embodied agent with postures and locomotion (S + L + P) showed higher social presence than the speech-only agent (S). Moreover, the same trend was also found in the copresence subscale of social presence in both studies. The S + L + P agents in our studies are embodied in the MR environment, walking in the same physical environment where users were. Study 2 further showed that the S + L copresence was significantly higher than for the S agent. Such spatial locomotion behavior demonstrates the agents' ability of contextual awareness, which has been found useful for fostering the copresence (Pimentel & Vinkers, 2021). Furthermore, consistent with the findings on postures reviewed by Wang and Ruiz (2021), the emotional cues conveyed through postures in the S + L + P condition enhanced the agents' interactivity and human-likeness, thereby further strengthening social presence.

Consistent with the social presence, we further found the embodied agent with postures (S + L + P) demonstrated higher social engagement than the embodied agent without postures (S + L). As noted by Oertel et al. (2020), engagement in human-agent interaction can be categorized into two types: 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 integrated the concepts of social engagement and task engagement, introducing the notion of social-task engagement, which arises in scenarios where human users and agents collaborate on a defined task.

In the two user studies presented in this paper, participants either played the Gobang game against the virtual agent (i.e., the virtual opponent in Study 1) or were assisted by it (i.e., the on-demand virtual assistant in Study 2), engaging with both the game logic and the agent's 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 studies. Additionally, social-task engagement was also present in both studies, as the player and the agent interacted collaboratively within the Gobang game, either by competing or providing assistance.

In both studies, social engagement was measured by the social engagement questionnaire, while behavioral measurements like GD proportion, task completion time, and task performance reflected task engagement. The EEG-engagement index could capture the overall social-task engagement as participants' cognitive activities were driven by brain activity, which the EEG signals could capture. As detailed in Table 2, the S + L + P agent demonstrated significantly higher average social engagement compared to the S + L agent in both studies, while there was no significant difference in task engagement as no significant differences were observed in any of the behavioral measurements across both studies.

For overall social-task engagement, the EEG-based engagement index at AF3 was significantly lower in the S + L + P condition than the S + L condition when the agent was an opponent in Study 1, whereas the EEG-based engagement index at P3 was significantly higher in the S + L + P condition compared to the S condition when the agent was assisting the user in Study 2. Both the EEG-based engagement index at AF3 and P3 reflected users' attention during the game (see discussions presented in Chang et al. (2024) and Section 6.5.2). However, the S + L + P opponent remained standing on the board at all times, whereas the S + L + P assistant stood at the side of the board and only walked onto it when players requested suggestions. In both studies, users' main task was to play the Gobang game while not interacting with the agents. Consequently, the S + L + P opponent's locomotion and postures in Study 1 might have distracted users from the task. In contrast, the S + L + P assistant's locomotion and postures in Study 2 likely enhanced overall social-task engagement, as the behaviors were not distracting but instead made the assistant more interactive and interesting, thereby increasing the overall engagement of the game. However, further research is needed to more fully understand the relationships among social engagement, task engagement, and overall social–task engagement in human–agent interactions, as well as to deepen our understanding of how physiological, behavioral, and subjective measures together capture engagement.

Responses to interview question **Q4** on agent role preference in Study 1 (Chang et al., 2024) revealed that participants with higher self-reported Gobang skill levels preferred the agent in an opponent role, while those with lower skill levels favored the assistant role. Similarly, results from Study 2 confirmed that novice players preferred interacting with an assistant. These findings suggest that users' individual preferences for the agent's role may significantly influence their perception of and interaction with the agent.

7.2. Design implications

By integrating the key findings in the presented two user studies, 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 engagement of a MiRA (based on both Studies).

DI2. Providing both locomotion and postures can enhance the social presence of a MiRA (based on both Studies).

DI3. The virtual assistant's social presence mediates rapport development (based on Study 2).

DI4. Although locomotion and postures in MiRAs can enhance social perception, they may also distract users, potentially leading to decreased task performance (based on Study 1).

DI5. 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 both studies).

DI6. 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 both studies).

DI7. The roles and behaviors of virtual agents should consider users' individual preferences to optimize interaction experiences (based on interviews in both Studies).

By integrating the design implications derived from our studies, virtual agents in mixed reality applications can significantly benefit domains such as education, therapy, and healthcare. For instance, virtual tutors in educational settings could use locomotion and postures (DI1, DI2) to foster greater social engagement and presence, thereby enhancing learning experiences. In therapy, socially aware agents could facilitate rapport development (DI3), enabling more effective emotional or behavioral interventions. However, care must be taken to balance social and task engagement (DI6) to avoid distractions that could hinder therapeutic or educational goals. In healthcare, virtual assistants could be personalized based on patient preferences (DI7), improving comfort and adherence during tasks like rehabilitation or mental health support. By adopting a holistic evaluation framework (DI5) that includes subjective, behavioral, and physiological measures, practitioners can ensure these agents deliver impactful, user-centered experiences while addressing potential tradeoffs such as the risk of reduced task performance (DI4).

7.3. Limitations

Our work has some limitations. First, mounting the Hololens headset on top of the EEG cap might disturb the movements of the EEG electrodes and thus result in poor signal data. This led to the loss of some EEG signals and might further explain why we could barely see a significant difference in EEG-based engagement at other electrodes. Future studies could address this limitation by utilizing advanced headsets that integrate EEG sensors with MR capabilities, such as OpenBCI's Galea[9]. Second, the sample size in both studies was small. With a larger sample, we anticipate stronger effect sizes for the current significant findings and expect the marginal significance observed between S + L and S in the EEG-based engagement index in Study 2 to reach statistical significance. Third, although both studies systematically tested the same three agent-behavior conditions (S, S + L, S + L + P), they were conducted with different participant populations and cannot be statistically compared. We therefore treat the cross-study synthesis as exploratory. Our findings also suggest that participants' preferences for agent role may depend on their skill level: novice players tended to prefer an assistant agent, whereas more experienced Gobang players preferred an opponent. A mixed-design study with a single participant pool, employing a 2 (role) $\times$ 3 (behavior) design, could more rigorously test these interactions and examine how skill level moderates the effect of agent role on user perceptions. Fourth, the speech of all agents in both studies was pre-scripted, limiting their interactivity and intelligence. Future improvements could leverage generative AI models, such as ChatGPT (Wu et al., 2023), to enhance these capabilities. Finally, 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.

8. Conclusion and future work

In this paper, we explored the role of locomotion and posture in MiRAs and their influence on perceived user engagement, social presence, and interaction quality in both competitive and cooperative contexts. Through two user studies involving a turn-based Gobang game, we demonstrated that MiRAs' non-verbal cues significantly shape users' social perceptions and engagement. Our findings highlight the nuanced effects of agent roles, with locomotion and posture enhancing social presence and engagement in competitive settings while fostering rapport but not trust in cooperative interactions.

We provided a holistic assessment of users' responses to MiRAs' non-verbal behaviors by integrating subjective, behavioral, and physiological measures, including EEG sensing. This multidimensional approach offers deeper insights into user-agent interactions and underscores the potential of EEG to capture cognitive and emotional responses beyond traditional measures.

Our contributions include empirical evidence of the impact of non-verbal cues in MiRAs across different contexts, an investigation of virtual assistants' behaviors in rapport building, and actionable design implications for enhancing human-agent interaction in mixed reality environments. These findings advance the understanding of MiRAs and offer valuable guidelines for designing engaging and socially effective virtual agents.

Future research could build on this work by examining the long-term effects of MiRAs' non-verbal cues, their influence across various application domains, and their interactions with other modalities such as gaze and facial expressions. Developing adaptive virtual agents that leverage physiological sensors like EEG to measure user engagement and attention, allowing for dynamic adjustments to task difficulty to maintain optimal engagement, is an exciting avenue for future exploration.

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://github.com/SSSxCCC/AlphaZero-In-Unity.
6. https://www.unicorn-bi.com/.
7. https://www.gtec.at/product/gnautilus-pro/.
8. https://mne.tools/stable/generated/mne.baseline.rescale.html.
9. https://galea.co/#home.

Author contributions

CRediT: **Zhuang Chang**: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft; **Kunal Gupta**: Conceptualization, Methodology, Supervision, Writing – review & editing; **Jiashuo Cao**: Conceptualization, Investigation, Methodology; **Huidong Bai**: Conceptualization, Methodology, Supervision; **Mark Billinghurst**: Project administration, Resources, Supervision, Writing – review & editing.

Disclosure statement

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

Ethics statement

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

References

Aguinis, H., Edwards, J. R., & Bradley, K. J. (2017). Improving our understanding of moderation and mediation in strategic management research. *Organizational Research Methods*, *20*(4), 665–685. https://doi.org/10.1177/1094428115627498

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. In*: Social Robotics: 13th International Conference, ICSR 2021, Singapore, Singapore, November 10–13, 2021, Proceedings 13* (pp. 671–682). Springer.

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).

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, *4*(7), 267–278. https://doi.org/10.1016/s1364-6613(00)01501-1

Andrist, S., Gleicher, M., & Mutlu, B. (2017). Looking coordinated: Bidirectional gaze mechanisms for collaborative interaction with virtual characters. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 2571–2582).

Aneja, D., Hoegen, R., McDuff, D., & Czerwinski, M. (2021). Understanding conversational and expressive style in a multimodal embodied conversational agent. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–10).

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. (2022). 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

Behrmann, M., Geng, J. J., & Shomstein, S. (2004). Parietal cortex and attention. *Current Opinion in Neurobiology*, *14*(2), 212–217. https://doi.org/10.1016/j.conb.2004.03.012

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. https://pubmed.ncbi.nlm.nih.gov/17547324/

Bickmore, T., & Cassell, J. (2001). Relational agents: A model and implementation of building user trust. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 396–403).

Biocca, F. (1997). The cyborg's dilemma: Progressive embodiment in virtual environments. *Journal of Computer-Mediated Communication*, *3*(2). https://doi.org/10.1111/j.1083-6101.1997.tb00070.x

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

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

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

Cerekovic, A., Aran, O., & Gatica-Perez, D. (2016). Rapport with virtual agents: What do human social cues and personality explain? *IEEE Transactions on Affective Computing*, *8*(3), 382–395. https://doi.org/10.1109/TAFFC.2016.2545650

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

Chang, Z., Cao, J., Gupta, K., Bai, H., & Billinghurst, M. (2024). Exploring the effects of mixed reality agents' locomotion and postures on social perception through a board game. *International Journal of Human–Computer Interaction*, 1–19. https://doi.org/10.1080/10447318.2024.2435694

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]. 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 1512–1515). IEEE.

Corrigan, L. J., Peters, C., Castellano, G., Papadopoulos, F., Jones, A., Bhargava, S., Janarthanam, S., Hastie, H., Deshmukh, A., & Aylett, R. (2013). *Social-task engagement: Striking a balance between the robot and the task. In: Embodied Commun* [Paper presentation]. Goals Intentions Workshop ICSR,. Vol. 13. pp. 1–7.

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of oz studies—Why and how. *Knowledge-Based Systems*, *6*(4), 258–266. https://doi.org/10.1016/0950-7051(93)90017-N

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

DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., Georgila, K., Gratch, J., Hartholt, A., Lhommet, M., et al. (2014). Simsensei kiosk: A virtual human interviewer for healthcare decision support. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems* (pp. 1061–1068).

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). Springer.

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

Ganesh, S., van Schie, H. T., de Lange, F. P., Thompson, E., & Wigboldus, D. H. (2012). How the human brain goes virtual: Distinct cortical regions of the person-processing network are involved in self-identification with virtual agents. *Cerebral Cortex (New York, N.Y.: 1991)*, *22*(7), 1577–1585. https://doi.org/10.1093/cercor/bhr227

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. S. (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

Grèzes, J., & de Gelder, B. (2008). Social perception: Understanding other people's intentions and emotions through their actions. In *Social cognition: Development, neuroscience and autism* (p. 5). Wiley-Blackwell.

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). IEEE. https://doi.org/10.1109/VR46266.2020.00099

Gupta, K., Zhang, Y., Gunasekaran, T. S., Krishna, N., Pai, Y. S., & Billinghurst, M. (2024). Caevr: Biosignals-driven context-aware empathy in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, *30*(5), 2671–2681. https://doi.org/10.1109/TVCG.2024.3372130

Harms, C., & Biocca, F. (2004). *Internal consistency and reliability of the networked minds measure of social presence* [Paper presentation]. Seventh Annual International Workshop: Presence. Vol. 2004. Universidad Politecnica de Valencia, Valencia, Spain.

Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (3rd ed.). Guilford Press.

He, Y., Pereira, A., & Kucherenko, T. (2022). Evaluating data-driven co-speech gestures of embodied conversational agents through real-time interaction. In *Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents* (pp. 1–8).

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. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1–13).

Huang, L., Morency, L.-P., & Gratch, J. (2011). Virtual rapport 2.0. In *Intelligent Virtual Agents: 10th International Conference, IVA 2011, Reykjavik, Iceland, September 15–17, 2011. Proceedings 11* (pp. 68–79). Springer.

Ishii, R., Ahuja, C., Nakano, Y. I., & Morency, L.-P. (2020). Impact of personality on nonverbal behavior generation. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents* (pp. 1–8).

Jeong, H., Yoo, J. H., Goh, M., & Song, H. (2023). Virtual agents in dtx: Focusing on usability and therapeutic effectiveness. *Electronics*, *13*(1), 14. https://doi.org/10.3390/electronics13010014

Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, *4*(1), 53–71. https://doi.org/10.1207/S15327566IJCE0401_04

Karim, S., Kang, J., & Girouard, A. (2023). Exploring rulebook accessibility and companionship in board games via voiced-based conversational agent alexa. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (pp. 2221–2232).

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. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (pp. 1–2).

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

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). IEEE. https://doi.org/10.1109/VR46266.2020.1581084624004

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

Kosonogov, V., Ntoumanis, I., Hajiyeva, G., & Jääskeläinen, I. (2024). The role of engagement and arousal in emotion regulation: An EEG study. *Experimental Brain Research*, *242*(1), 179–193. https://doi.org/10.1007/s00221-023-06741-3

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). IEEE.

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., Hollerer, T., & Welch, G., (2018a). 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. (2018b). The physical-virtual table: Exploring the effects of a virtual human's physical influence on social interaction. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (pp. 1–11).

Lee, M., Norouzi, N., Bruder, G., Wisniewski, P. J., & Welch, G. F. (2019). 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). IEEE.

Leite, I., Mascarenhas, S., Pereira, A., Martinho, C., Prada, R., & Paiva, A. (2010). *"Why can't we be friends?" An empathic game companion for long-term interaction* [Paper presentation]. International Conference on Intelligent Virtual Agents (pp. 315–321). Springer.

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. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents* (pp. 191–196).

Li, X., Zhang, W., Chen, J., & Wu, L. (2022). *Gobang game algorithm based on reinforcement learning* [Paper presentation]. Cognitive Systems and Information Processing: 6th International Conference, ICCSIP 2021 (pp. 463–475), Suzhou, China. Cairangdanghzhou, November 20–21, 2021. Revised Selected Papers 6. Springer.

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

McCleery, A., Horan, W. P., & Green, M. F. (2014). Social cognition during the early phase of schizophrenia. In *Social cognition and metacognition in schizophrenia* (pp. 49–67). Elsevier.

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). IEEE.

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

Morel, B., & Ach, L. (2011). *Attention-aware intelligent embodied agents*. Cambridge University Press.

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). Springer.

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. *Artificial intelligence in IoT* (pp. 1–24).

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. In *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

Pereira, A., Prada, R., & Paiva, A. (2014). Social presence and artificial opponents. *Entertaining the Whole World*, 115–135.

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). IEEE. https://doi.org/10.1109/CoG52621.2021.9618990

Pimentel, D., & Vinkers, C. (2021). Copresence with virtual humans in mixed reality: The impact of contextual responsiveness on social perceptions. *Frontiers in Robotics and AI*, *8*, 634520. https://doi.org/10.3389/frobt.2021.634520

Plopski, A., Hirzle, T., Norouzi, N., Qian, L., Bruder, G., & Langlotz, T. (2022). The eye in extended reality: A survey on gaze interaction and eye tracking in head-worn extended reality. *ACM Computing Surveys*, *55*(3), 1–39. https://doi.org/10.1145/3491207

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. (2019a). 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

Randhavane, T., Bera, A., Kapsaskis, K., Sheth, R., Gray, K., & Manocha, D. (2019b). *Eva: Generating emotional behavior of virtual agents using expressive features of gait and gaze* [Paper presentation]. ACM Symposium on Applied Perception 2019 (pp. 1–10).

Rheu, M., Shin, J. Y., Peng, W., & Huh-Yoo, J. (2021). Systematic review: Trust-building factors and implications for conversational agent design. *International Journal of Human–Computer Interaction*, *37*(1), 81–96. https://doi.org/10.1080/10447318.2020.1807710

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). IEEE.

Rushworth, M. F., Krams, M., & Passingham, R. E. (2001). The attentional role of the left parietal cortex: The distinct lateralization and localization of motor attention in the human brain. *Journal of Cognitive Neuroscience*, *13*(5), 698–710. https://doi.org/10.1162/089892901750363244

Rushworth, M. F., Nixon, P. D., Renowden, S., Wade, D. T., & Passingham, R. E. (1997). The left parietal cortex and motor attention. *Neuropsychologia*, *35*(9), 1261–1273. https://doi.org/10.1016/s0028-3932(97)00050-x

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. IUI '04 (pp. 78–84). Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/964456.964458

Steven, J. (2014). *An introduction to event related potential technique* (2nd ed.). The MIT Press.

Sun, N. (2023, November 14). *Delegation to virtual agents in critical scenarios: Influencing factors and immersive settings* [Ph.D. thesis]. Unilu - University of Luxembourg [FSTM].
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). IEEE.
Szafir, D., & Mutlu, B. (2012). Pay attention! designing adaptive agents that monitor and improve user engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 11–20).
Tan, S.-M., & Liew, T. W. (2020). Designing embodied virtual agents as product specialists in a multi-product category e-commerce: The roles of source credibility and social presence. *International Journal of Human–Computer Interaction*, *36*(12), 1136–1149. https://doi.org/10.1080/10447318.2020.1722399
Torre, R., Fua, P., Balcisoy, S., Ponder, M., & Thalmann, D. (2000, June 1–2). Interaction between real and virtual humans: Playing checkers. In *Virtual Environments 2000: Proceedings of the Eurographics Workshop in Amsterdam, The Netherlands* (pp. 23–32). Springer.
Wang, H., Preuss, M., & Plaat, A. (2021). Adaptive warm-start mcts in alphazero-like deep reinforcement learning. In *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). Springer.
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. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–12).
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
Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q.-L., & Tang, Y. (2023). A brief overview of chatgpt: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica*, *10*(5), 1122–1136. https://doi.org/10.1109/JAS.2023.123618
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). IEEE.
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.

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.

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.

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.