



OPEN ACCESS

EDITED BY
Jonathan W. Kelly,
Iowa State University, United States

REVIEWED BY
Xiaoli Chen,
Zhejiang University, China
Tulin Gunduz Cengiz,
Uludağ University, Turkey

*CORRESPONDENCE
Zhuang Chang,
zcha621@aucklanduni.ac.nz

SPECIALTY SECTION
This article was submitted to Virtual Reality and Human Behaviour, a section of the journal Frontiers in Virtual Reality

RECEIVED 15 July 2022
ACCEPTED 11 November 2022
PUBLISHED 29 November 2022

CITATION
Chang Z, Bai H, Zhang L, Gupta K, He W and Billingham M (2022), The impact of virtual agents' multimodal communication on brain activity and cognitive load in Virtual Reality. *Front. Virtual Real.* 3:995090. doi: 10.3389/frvir.2022.995090

COPYRIGHT
© 2022 Chang, Bai, Zhang, Gupta, He and Billingham. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The impact of virtual agents' multimodal communication on brain activity and cognitive load in Virtual Reality

Zhuang Chang^{1*}, Huidong Bai¹, Li Zhang², Kunal Gupta¹, Weiping He² and Mark Billingham¹

¹Empathic Computing Laboratory, Auckland Bioengineering Institute, The University of Auckland, Auckland, New Zealand, ²Cyber-Physical Interaction Laboratory, Northwestern Polytechnical University, Xi'an, China

Related research has shown that collaborating with Intelligent Virtual Agents (IVAs) embodied in Augmented Reality (AR) or Virtual Reality (VR) can improve task performance and reduce task load. Human cognition and behaviors are controlled by brain activities, which can be captured and reflected by Electroencephalogram (EEG) signals. However, little research has been done to understand users' cognition and behaviors using EEG while interacting with IVAs embodied in AR and VR environments. In this paper, we investigate the impact of the virtual agent's multimodal communication in VR on users' EEG signals as measured by alpha band power. We develop a desert survival game where the participants make decisions collaboratively with the virtual agent in VR. We evaluate three different communication methods based on a within-subject pilot study: 1) a Voice-only Agent, 2) an Embodied Agent with speech and gaze, and 3) a Gestural Agent with a gesture pointing at the object while talking about it. No significant difference was found in the EEG alpha band power. However, the alpha band ERD/ERS calculated around the moment when the virtual agent started speaking indicated providing a virtual body for the sudden speech could avoid the abrupt attentional demand when the agent started speaking. Moreover, a sudden gesture coupled with the speech induced more attentional demands, even though the speech was matched with the virtual body. This work is the first to explore the impact of IVAs' interaction methods in VR on users' brain activity, and our findings contribute to the IVAs interaction design.

KEYWORDS

intelligent virtual agents, Virtual Reality, multimodal communication, EEG, cognitive load

1 Introduction

People mainly communicate with each other using verbal and non-verbal cues, such as eye gaze, facial expressions, and hand gestures, to solve problems collaboratively. Previous research shows that Intelligent Virtual Agents (IVAs) embodied with human-like characteristics (e.g., human appearance, natural language, and gestures) in Virtual Reality (VR) or Augmented Reality (AR) could be treated like an actual human in human-agent collaboration (Hanton et al., 2016; Li et al., 2018). Thus, it would be interesting to explore efficient interaction modes in IVAs for human-agent collaboration in VR and AR scenes.

Neuroscience research has shown that different brain areas control human brain activity. Electroencephalogram (EEG) has been shown effective in recording brain activity (Oostenveld and Praamstra, 2001), and thus has been widely used as an objective measurement in human computer interaction research (Lécuyer et al., 2008; Fink et al., 2018; Gerry et al., 2018). Our research is particularly interested in how IVA representation affects brain activity.

Prior research showed that IVAs embodied in AR could be helpful in improving human-agent collaboration by reducing task load (Kim et al., 2020), enhancing a sense of social presence (Wang et al., 2019), and improving task performance (de Melo et al., 2020). However, there has been little research on whether interacting with IVAs in VR/AR can affect users' physiological states by measuring their brain activity.

In this paper, we conduct a pilot study to explore how the IVAs' communication methods influence users' EEG signals. Inspired by the AR desert survival game (Kim et al., 2020), we developed a VR version where participants worked collaboratively with IVAs to rank 15 desert survival items in three conditions: 1) working with a voice-only agent providing verbal cues only, 2) working with an embodied agent communicating through both verbal cues and eye gaze, and 3) working with a gestural agent expressing information by looking and pointing at the object while talking about it. We focus our research on understanding the effects of visual cues exhibited by IVA's behaviors on users' brain activity and subjective user experiences.

Compared to prior work, our research makes two key contributions. Firstly, it is the first research that explores the effects of IVA's interaction approaches on users' brain activity for a collaborative task in VR. Secondly, we present a pilot study to examine the relationship between IVAs' interaction methods and the user's EEG signals.

2 Related work

We focus our review of previous work on three categories. First, we highlight the importance of multimodal communication of embodied IVAs in AR and VR. Second, we present the theory of cognitive load. Third, we introduce the current methods of using EEG to measure cognitive load.

2.1 Multimodal communication of embodied intelligent virtual agents in Augmented reality and Virtual reality

Designing anthropomorphic IVAs to communicate with users using both verbal and non-verbal cues in VR and AR has gained much attention (Holz et al., 2011; Norouzi et al., 2019; Norouzi et al., 2020). Visual embodiment and social behaviors like the agent's gesture and locomotion could improve perceived social presence in both AR (Kim et al., 2018) and VR (Ye et al., 2021). Li et al. (2018) investigated how embodiment and postures influence human-agent interaction in Mixed Reality (MR), in which they found people treated virtual humans similar to real persons. Such humanoid IVAs in AR have been proven to be useful for helping people with real-world tasks (Ramchurn et al., 2016; Haesler et al., 2018). In the context of VR training, Kevin et al. (2018) explored interaction with a gaze-aware virtual teacher where the virtual agent reacted differently to students' eye gaze in a variety of situations. They found this type of interaction resulted in an overall better experience, including social presence, rapport, and engagement. These exemplify the importance of visual cues coupled with voice information for the virtual agent interaction design.

However, it is controversial on designing visual cues for virtual agent interaction. Some research has shown IVA visual embodiment causes distractions, especially when the task difficulty increases (Miller et al., 2019). And the realism level of the agent's appearance may also induce uncanny valley effects (Reinhardt et al., 2020). Therefore, understanding the effects of virtual agent's non-verbal cues (Wang and Ruiz, 2021) on user experience and cognition is important.

Current research involving user experience on interacting with IVA embodied in VR and AR mainly use questionnaires and system logged user behaviors as measurements (Norouzi et al., 2018). For example, Wang et al. (2019) conducted a user study to explore how the visual representation of virtual agent affects user perceptions and behavior. Participants were asked to solve hidden object puzzles with four types of virtual agents in AR: voice-only, non-human, full-size embodied and miniature embodied agent. They logged user-agent interactions such as user and agent utterance, times the user gazed at the agent and so on. Except for the logged interaction behaviors, they also collected questionnaires on user preference and ratings on helpfulness, presence, reliability, trust, distraction, and realism. Similarly, Kim et al. (2020) designed a miniature IVA in AR for assisting users to make decisions on ranking items for the desert survival task through speaking, body postures, and locomotion. In their study, participants performed the task under three conditions: 1) working alone, 2) working with a voice assistant, and 3) working with an embodied assistant. The embodied virtual agent was most helpful for reducing task load, improving task performance, and the feeling of social presence and social richness.

As a common knowledge, user behaviors and perceptions are controlled by brains. Prior research has demonstrated the effects of interacting with humanoid robots (Suzuki et al., 2015) and IVAs embodied in two-dimensional screens (Mustafa et al., 2017) can be reflected in users' brain activity. However, little research has been done to investigate the impact of interacting with IVA in AR and VR (three-dimensional virtual world) on users' brain activity. To fill this gap, we ran a pilot study to explore the impact of agent interaction methods in VR on user's brain activity as measured by EEG signals. Similar to (Kim et al., 2020), we developed a VR version desert survival game and asked participants to play this game in VR with the help of a virtual agent. Except for measuring EEG signals, we also used questionnaires and system logged data to capture users' cognitive load, feeling of being together with the virtual agent (copresence) and task performance. We expected the results of EEG signals, subjective questionnaires and system logged data can have identical or complementary results in describing the user perceptions during the interaction.

2.2 Cognitive load theory

Cognitive load theory (CLT) was originally developed to address the interaction between information architecture and cognitive architecture in the fields of educational psychology and instructional design (Sweller, 1988; Sweller et al., 1998; Paas et al., 2004). The CLT focuses on the learning process of complex cognitive tasks, where learners are often faced with overwhelming amount of interactive information elements and their interactions required to be processed parallelly before meaningful understanding can commence (Paas et al., 2004). This complex information architecture was thought to have driven the evolution of cognitive architecture which consists of long-term memory (LTM) and working memory (WM) (Paas et al., 2003a). The effectively unlimited LTM contains vast number of schemas which are cognitive constructs that incorporate multiple elements of previously acquired information into a single element with a specific function. For example, the rules of matrix multiplication integrate the elements of basic numerical multiplication and addition. In contrast, the WM, in which all conscious cognitive activities occur, is very limited in both capacity and duration.

Based on the interaction between LTM and WM, the CLT identifies three types of cognitive load: intrinsic, extraneous and germane (Paas et al., 2003a; Paas et al., 2003b). The intrinsic cognitive load is imposed by the number of information elements and their interactivity whereas the extraneous and germane cognitive load is imposed by the manner in which the information is presented and the learning activities required of the learners. For example, adding two numbers less than ten causes much less intrinsic cognitive load than multiplying

two three-digit numbers. The extraneous cognitive load is generated by the manner in which the information is presented. For example, learning the structure of three-dimensional (3D) geometry with 3D models induces less extraneous cognitive load than learning with two-dimensional (2D) geometry views. The germane cognitive load is the load contributes to the construction and automation of schemas. For example, the cognitive load in understanding the operational rules of matrix multiplication is germane cognitive load as it dedicates to the schema construction of the matrix multiplication.

The CLT has been integrated and developed in the Human Computer Interaction (HCI) research field (Hollender et al., 2010). In the HCI context, the CLT is concerned with the cognitive load in completing HCI tasks. Therefore, measuring cognitive load or work load in HCI research is popular. For example, Jing et al. (2021), adopted SMEQ questionnaire (Sauro and Dumas, 2009) to measure workload in evaluating three bi-directional collaborative gaze visualizations with three levels of gaze behaviors for co-located collaboration. The cognitive load in their research task is extraneous cognitive load as the researchers manipulated the manner of communicating cues. Dey et al. (2019) used the N-back game (participants have to allocate WM resources to memorize and recall the correct information appeared N-rounds ago) to induce intrinsic cognitive load and adopted EEG to measure that type of cognitive load.

This work is dedicated to explore the effects of IVA's representations on human cognitive load and brain activity. We designed a VR version desert survival game, in which a virtual agent was integrated. This game involved both intrinsic cognitive load (making decisions on the order of desert survival items) and extraneous cognitive load (interacting with different types of virtual agent). We also adopted both EEG and questionnaires to measure cognitive load.

2.3 Electroencephalogram and cognitive load

Measuring cognitive workload from EEG signals has been researched for a long time (Klimesch, 1999; Antonenko et al., 2010). EEG signals can be decomposed into four components in the frequency domain, including Beta (> 13 Hz), Alpha (8–13 Hz), Theta (4–8 Hz), and Delta (0.5–4 Hz) bandwidths (Teplan, 2002). The power spectrums of these bands obtained from different brain locations can reflect different types of cognitive load (Kumar and Kumar, 2016). For example, delta power has been showed to increase as the intrinsic cognitive load induced by mental calculation and short-term memory tasks increased (Harmony et al., 1996). Much research has reported frontal theta power tended to increase with the increasing task difficulty whereas the alpha power in parietal lobe tended to decrease (Gevins and Smith, 2003; Holm et al., 2009).

Compared with using beta and delta band to measure cognitive load, using theta and alpha are better researched (Klimesch, 1999; Antonenko et al., 2010). However, theta band power is mainly prominent at frontal lobe (Klimesch, 1999; Freeman, 2002; Antonenko et al., 2010) while the alpha band power can be prominent at frontal lobe, parietal lobe and occipital lobe (Lang et al., 1988; Klimesch, 1999; Fink and Benedek, 2014). Besides, alpha rhythm is the dominant rhythm in normal human EEG and is most extensively studied (Teplan, 2002; Antonenko et al., 2010). Therefore, in this work we mainly focus on the alpha band frequency.

Using alpha band power to measure cognitive load induced by interacting VR scenes has attracted more and more attention (Zhang et al., 2017; Gerry et al., 2018; Gupta et al., 2019; Gupta et al., 2020). Dey et al. (2019) explored a cognitively adaptive VR training system based on real-time measurement of task related alpha power. They used the N-back game to induce cognitive load and calculated alpha power with data collected from prefrontal cortex, parietal lobe and occipital lobe. The task complexity adapted according to the rule of “lower alpha power, higher task difficulty”. Similarly, Gupta et al. (2020), used alpha power calculated with data collected from the same brain areas to measure the cognitive load of interacting with a virtual agent. We followed these two research and collected EEG data from prefrontal cortex, parietal lobe and occipital lobe as well.

Except for measuring cognitive load using spectrum power directly, the Event Related (De-)Synchronization (ERD/ERS) was proposed to calculate the changes in EEG signals (e.g., alpha band power) caused by a certain event (Antonenko et al., 2010). The positive value of ERD/ERS in alpha band represents a decrease in alpha power (ERD) and the negative value means an increase in alpha power (ERS) (Klimesch et al., 1992). Since ERD/ERS could reflect shifts in alpha band power between pre-stimulus and post-stimulus, it could be used to measure cognitive load as well. For example, Stipacek et al. (2003) varied the cognitive load of memory tasks to test the sensitivity of alpha band ERD in reflecting levels of cognitive load. They reported that ERD in the upper alpha band grew linearly with increasing cognitive load.

In this paper, we were interested in both alpha band power and alpha ERD/ERS in participants' EEG signals by varying the virtual agent's interaction methods. We expected the alpha band power could reflect the general cognitive load or level of neural relaxation during interacting with the virtual agent. And ERD/ERS would be used to reflect the sudden visual/audio changes in cognitive load or attentional demands at a specific event moment.

3 Materials and methods

In this section, we present the IVAs, the VR environment, and the implementation of our prototype system. We developed three different IVA types for the experiment: a voice-only agent,

an embodied agent with head gaze and eye gaze, and a gestural agent with a pointing gesture.

3.1 Intelligent virtual agents

Voice-only agent: Like most commercial assistants, such as Amazon Alexa, Microsoft Cortana and Apple Siri, our voice agent has no visual representation while only providing help to users while being asked through speech input. For example, the voice agent captures customized keywords in users' voices, and then responds to the recognized keywords according to predefined configuration files by incorporating IBM Watson Speech to Text¹ and Text to Speech services² into the agent communication system.

Embodied agent: A full body of a 3D male character was generated through Ready Player Me³ for the embodied agent. The character's blend shapes contained fifteen Lipsync visemes⁴ which were used for synchronizing the agent's lip shapes with the agent's voices. Tiny body movement animation was applied to the character, and a live shadow was projected on the ground in front of the agent to make it look realistic. Before talking about an object in VR with the user, the embodied agent will look at it with the upper limb and head slightly turning toward it. In short, the Embodied Agent is an agent embodied with eye gaze and head gaze.

Gestural agent: Based on the embodied agent, we further developed the agent to be capable of pointing at an object in VR with the left hand while commenting on it. The hand was put back to the rest position once the agent speaking stopped. In short, the Gestural Agent is an agent embodied with pointing gesture, head gaze and eye gaze.

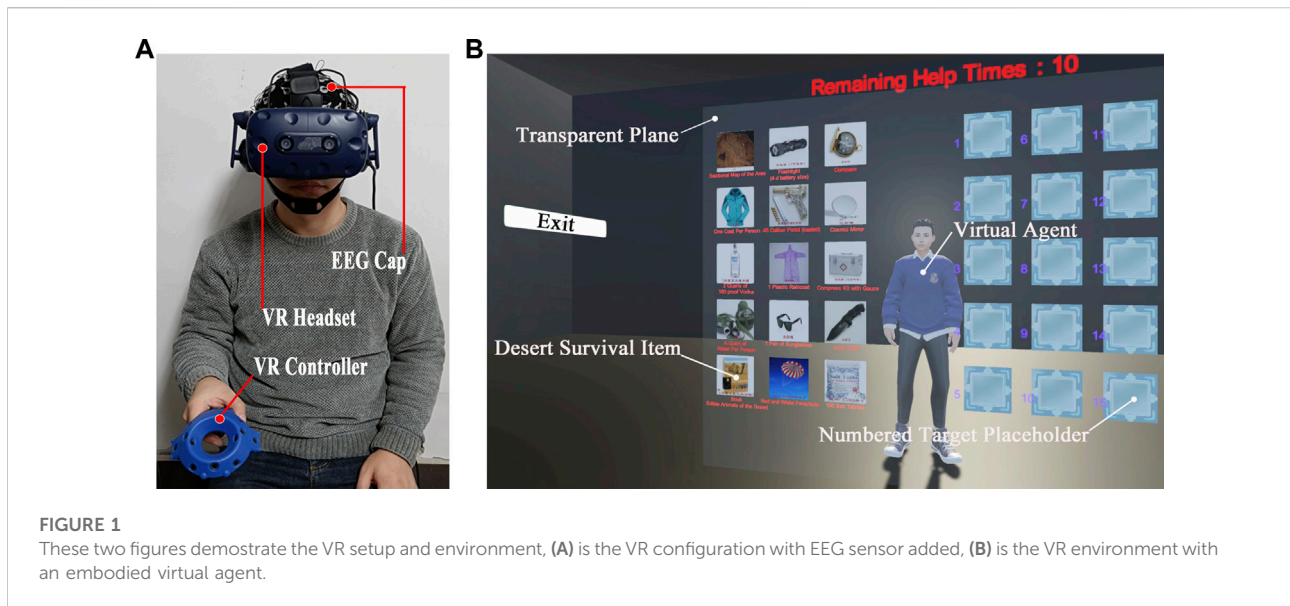
Since the agent speech system relays on IBM cloud service, there is around 2~3 s delay between user request and agent speech response due to the internet transfer. This time delay is allowed in our design for two reasons. First, this delay can be useful to simulate the thinking process like a real human normally does before given suggestions to partners in collaborative decision-making scenarios. Second, participants will rest shortly with least body movement during this period to wait for the agent's suggestions. EEG signals captured during this period can be used as baseline interval data for calculating ERD/ERS (see Section 8 for more details).

¹ <https://cloud.ibm.com/catalog/services/speech-to-text>

² <https://cloud.ibm.com/catalog/services/text-to-speech>

³ <https://readyplayer.me/avatar>

⁴ <https://developer.oculus.com/documentation/unity/audio-ovrlipsync-viseme-reference/>



3.2 Virtual reality environment

To reduce distractions from the surrounding environment, we constructed a simple house with bare walls in VR. At the house center, there was a transparent plane, with a 5×3 pattern with 15 desert survival items and another similar pattern with 15 numbered target placeholders (see Figure 1). For the embodied or gestural agents, a virtual character would stand behind the transparent plane with its entire body being seen through the gap between the two patterns. A label above the transparent plane would indicate the remaining help requests with a default value of 10 in all conditions.

3.3 Prototype implementation

For hardware, we used a Neuracle⁵ NSW364 EEG cap with 64 wet channels working at a 1000 Hz sampling rate, and the default reference electrode was CPz. We also used an HTC VIVE Pro Eye VR headset to display the VR environment running on a laptop powered by the Intel Core i7 8750 CPU and the NVIDIA GTX 1070 GPU. For software, we used the Neuracle EEGRecorder V2.0.1 for recording EEG signals and the Unity game engine for the game logic implementation and the VR application rendering.

⁵ <https://www.crunchbase.com/organization/neuracle>

4 Pilot study

We conduct a pilot study to evaluate the impact of the IVAs interaction method on users' EEG signals and experience. We recruited 11 right-handed participants (10 male and 1 female) with their ages ranging from 22 to 37 years old ($M = 25.4$, $SD = 4.3$). All of them had some experience with VR interfaces, and none of them reported a history of neurological disorders or had experience with playing the desert survival game.

The experiment used a within-subject design, and the independent variable was the types of IVA forming three conditions: 1) Voice-only Agent (baseline) communicating through speech, 2) Embodied Agent interacting with speech, head gaze and eye gaze, and 3) Gestural Agent embodied with a pointing gesture based on the Embodied Agent (see Figure 2). We used a Balanced Latin Square to balance the learning effect. In our study, we would like to explore the following research question:

How does the IVA communication approach affect the users' brain activity?

Our research hypotheses are:

1) **H1** Compared to the baseline, working with visually embodied IVAs would improve collaboration in VR (as measured by task performance, copresence and workload)

2) **H2** EEG signals will be significantly different among the three agent interaction conditions (as measured by EEG alpha band mean power and alpha ERD/ERS).

We made the H1 for two reasons. First, prior research has shown the visually embodied virtual agent in AR could improve task performance and social presence and reduce task load in a similar desert survival game (Kim et al., 2020). Thus, we expected some of these effects carry on in our experiment. Second, we expected the results of copresence and workload in H1 could

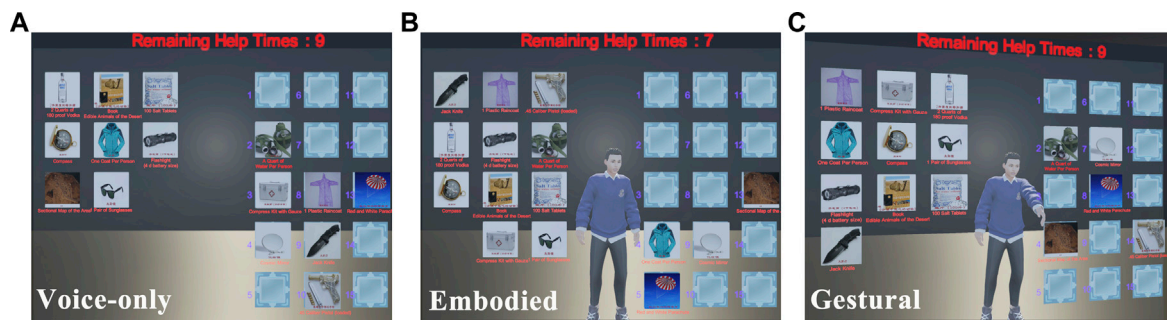


FIGURE 2

These three figures show the three agent types, (A) is the Voice-only Agent, (B) is the Embodied Agent, (C) is the Gestural Agent.

support the analysis of EEG results examined in the H2. As we described in Section 2.3, the EEG alpha suppression could reflect cognitive load or wakeful relaxation. Combining the results of H1, we hoped the results of H2 could answer the research question we proposed.

5 Experiment set-up

The experiment was conducted in an isolated space with minimum radio interference with three computers for running our VR program, collecting questionnaires, and recording EEG signals, respectively. Participants wearing both EEG cap and HTC VIVE headset (see Figure 1) were seated 1 m away from the real-world table and around 2 m away in front of the virtual transparent plane, opposite to the virtual human positioned around 0.8 m away behind the transparent virtual plane.

6 Experimental task

We designed a VR version of the desert survival game. The main game logic was identical to the reference, except for how agents provide help and the scoring method. Participants' goal was to rank the 15 desert survival items by dragging them from the left pattern to the right side numbered target placeholders. At the beginning of each session, the initial positions of the 15 items in the left-hand pattern were randomized while the numbers of the target placeholder always started from 1 to 15, representing the most critical position to the least. Participants used an HTC VIVE controller with a virtual ray attached for object selection. Items could be attached to the ray's end and moved when participants held the controller trigger. Once the selected item was placed on top of the target place, it would be automatically attached to that placeholder when the trigger was released.

The virtual agent in each condition provided help for up to ten times. When at least one item was placed on one of the target

placeholders, participants could ask for help by saying "help". After 2~3 s time delay caused by the internet transfer, the virtual agent would then suggested one item to move, helping participants achieve a better score. For both embodied condition and gestural condition, the virtual agent would look at the target placeholder area once the participants requested for help. The gestural agent's pointing gesture is coupled with the speech.

The virtual agents' speech was made up of positive or negative reasons and weak or strong movement suggestions. For example, the positive reason with the weak suggestion for the coat would be "the coat could slow down the dehydration process, I suggest moving it up a bit". In contrast, the negative reason with the strong suggestion for the vodka would be "the vodka is almost useless except for starting a fire, I think it should be placed at the end".

The task score was recalculated every time the items in placeholders were changed. We gave 15 points to the most important items and 1 point to the least important ones. We summed the points of 15 items getting 120 as the total score, and then calculated the participant's final score by reducing the sum of absolute difference values between the expert answer (Kim et al., 2020) and the participants' choice. Once the participants complete the game, the final task score will be saved in each session.

7 Experimental procedure

Initially, participants entered the testing room to fill out a consent form with an opportunity to ask questions about the study. Once they agreed to participate in the experiment, a video clip was played to demonstrate an extreme desert environment and fifteen objects for the desert survival game. The investigator explained the task goal and the three different experimental conditions in detail. Afterward, participants sat in front of a table with two computers for running the VR scenes and recording EEG data. They wore the HTC VIVE headset to practice dragging and placing virtual pictures of desert survival items to different given placeholders using a controller shooting a virtual ray on the object.

Once they felt confident interacting with the virtual pictures, the investigator helped them take off the VR headset and put on the EEG cap. Then we injected gel into nine interested electrodes (GND, REF, Fp1, Fp2, P3, P4, Pz, O1, and O2) to reduce the impedance level below 40 K Ω (Luck, 2014; Gupta et al., 2020). After filling the gel, we double-checked each channel's state by observing signals on the EEG data capture computer while participants were told to blink, clench their jaw, close their eyes for several seconds and open suddenly.

Afterward, participants waited for 5 minutes allowing the gel to settle in. We then helped the participants put on the VR headset and start the main task. At the end of each session, participants were asked to fill out the Copresence questionnaire (Pimentel and Vinkers, 2021), the Subjective Mental Effort Questionnaire (SMEQ) (Zijlstra and Van Doorn, 1985; Sauro and Dumas, 2009) and the NASA Task Load Index (NASA-TLX) questionnaire (Hart and Staveland, 1988). We also interviewed the participants with open-ended questions to understand their experiences and feelings on the IVA. Each participant took three trials in total and the experiment mostly costs about 90 minutes.

8 Measurement

We collected raw EEG data from the prefrontal cortex (Fp1, Fp2), parietal lobe (P3, P4, and Pz), and occipital lobe (O1, O2) based on the 10–20 system (Oostenveld and Praamstra, 2001), because these brain areas were proved to be about decision-making (Miller and Cohen, 2001), attentional demands (Klimesch, 1997), and vision processing (Malach et al., 1995), respectively. We expected to measure the intrinsic cognitive load caused by making decisions from the frontal cortex and the extraneous cognitive load from the parietal and occipital lobe. Due to technical reasons, one participant's EEG data was not recorded correctly and thus excluded from the EEG processing.

To remove EEG artifacts induced by wires, eye blinking, swallow, the slight shift in the VR headset's position, etc., we manually checked each participant's EEG signals for each condition and removed bad channels. Then we applied a bandpass filter (1~40 Hz) followed by an independent component analysis (ICA). After the preprocessing, we computed alpha band power and ERD/ERS for each channel with the preprocessed data.

For calculating EEG alpha power, we only focused on the EEG signals when agent was speaking. We made this decision for three reasons. First, this research mainly focus on the effect of agent's interaction methods on users' brain activity. The agent-speaking time window is when the participants and virtual agent were actually interacting. Second, the average agent speech length was 5.1 s which is long enough for containing meaningful information. Much EEG research focused on milliseconds data (Luck, 2014). Third, the remaining EEG signals might contain interference caused

by body movements for adjusting the object orders with VR controllers.

Furthermore, we chunked each agent-speaking time windows into 1 s epochs and used each epoch data to calculate alpha band power. With regard to a single channel, we averaged the alpha power for each agent-speaking window. Before performing the statistical tests on alpha band power, we grouped data from related channels into prefrontal (Fp1, Fp2), parietal (P3, P4, and Pz) and occipital (O1, O2) groups.

For calculating ERD/ERS, we used the following formula (Stipacek et al., 2003):

$$\begin{aligned} &\text{ERD/ERS in alpha band} \\ &= \frac{\text{baseline interval alpha band power} - \text{test interval alpha band power}}{\text{baseline interval alpha band power}} * 100\% \end{aligned} \quad (1)$$

We treated the time when agent started speaking as stimulus onset. We selected -1 s pre-stimulus as baseline time interval and 1 s post-stimulus as test interval (see Figure 3). The ERD/ERS was only calculated with parietal and occipital data as we did not expect any decision-making cognitive load at the moment when the virtual agent started speaking. Besides, the sudden visual/audio change would cost demands in attentional and visual processing. Therefore, we expected the parietal and occipital lobe could capture this type of cognitive load.

Similar to Kim et al. (2020), we also collected data on task performance, subjective ratings on task load and feeling of copresence to understand whether the variety of the virtual agent in our research would influence the collaboration. We used the task score logged from each session to measure the task performance. We also collected data using the Copresence questionnaire (Pimentel and Vinkers, 2021), SMEQ (Sauro and Dumas, 2009) and NASA-TLX questionnaire (Hart and Staveland, 1988) to investigate subjective user experience on the feelings of being together with virtual humans and the workload of the collaborative task for each condition. Since the SMEQ has only one question on the overall feeling of cognitive load while the NASA-TLX has six items focusing on both mental and physical efforts to finish the task, we used both to obtain comprehensive understanding on the cognitive load of our task.

9 Results

9.1 Electroencephalogram signals

9.1.1 Alpha band mean power

The Shapiro-Wilk normality test on the EEG alpha band mean power indicated it was not following a normal distribution. The Friedman test showed no significance difference among the

three agent conditions in prefrontal, parietal and occipital brain areas (see Figure 4).

9.1.2 Alpha band the event related (De-)synchronization

The Shapiro-Wilk tests on alpha band ERD/ERS from the interested brain areas revealed that the normality assumption had been violated. Friedman tests revealed both the parietal ($\chi^2(2) = 15.48, p < 0.001$) and occipital ($\chi^2(2) = 13.60, p = 0.001$) group showed significant difference among the three agent conditions. However, the follow-up Kendall's W tests revealed both effect sizes were small ($W = 0.02$ for parietal lobe and $W = 0.06$ for occipital lobe).

As shown in Figure 5, the parietal alpha ERD/ERS value was negative (ERS) of the Embodied Agent condition whereas the values of Gestural Agent condition and Voice-only Agent condition were both positive (ERD). For the occipital lobe, both the ERD/ERS values of the Embodied and Gestural conditions were negative (ERS) while that of the Gestural condition was positive (ERD).

Wilcoxon signed rank test with Boferroni correction was adopted to perform pairwise comparisons between agent conditions in both parietal and occipital brain areas. For the parietal lobe, we found alpha band ERD values of the Gestural ($Z = -3.29, p = 0.001$) and Voice-only ($Z = -3.09, p = 0.002$) condition was significantly different from the ERS of the Embodied condition. No significant difference was

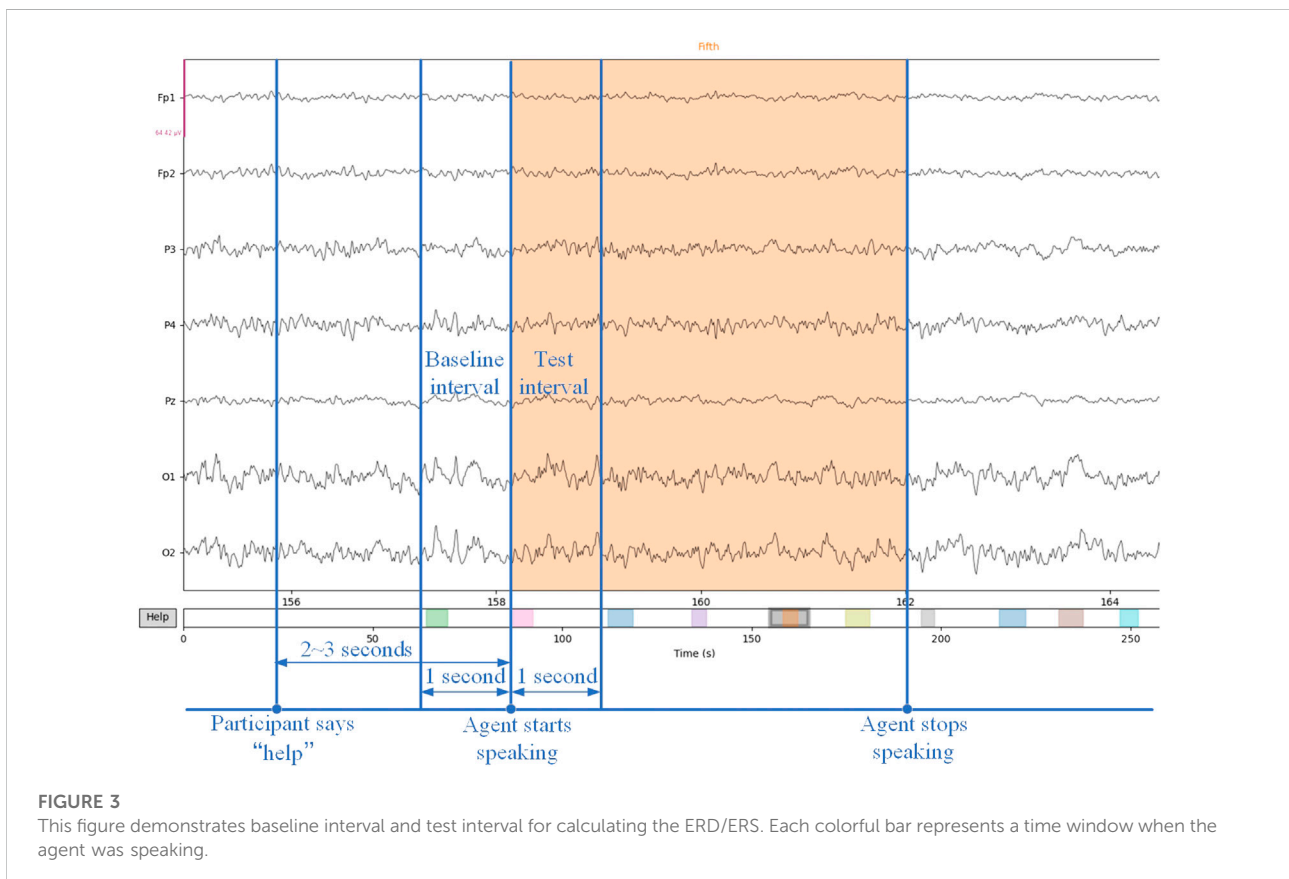
found between the ERD values of Gestural and Voice-only agent condition. For the occipital lobe, we only found significant difference between the ERS of the Embodied condition and the ERD of the Gestural condition ($Z = -3.29, p = 0.001$).

9.2 Task performance

One-way repeated measures ANOVA found no significant difference ($F(2, 20) = 1.05, p = 0.37, \eta^2 = 0.08$) among the three conditions. Descriptive statistics showed the task score of the Gestural Agent condition ($M = 107.09, SD = 8.19$) was highest, followed by the condition of Embodied Agent condition ($M = 106.36, SD = 11.67$) and the Voice-only Agent condition ($M = 100.45, SD = 11.54$).

9.3 Subjective results

A Shapiro-Wilk normality test showed that the copresence data of all three conditions violated the normality assumption. A significant difference in copresence over the three conditions was found using the Friedman test ($\chi^2(2) = 19.86, p < 0.0001$). Large effect size ($W = 0.90$) was found by the following Kendall's W test. Post-hoc Wilcoxon signed-rank test with Bonferroni correction revealed that all the pairwise differences were significantly different ($p < 0.001$) (see Figure 6).



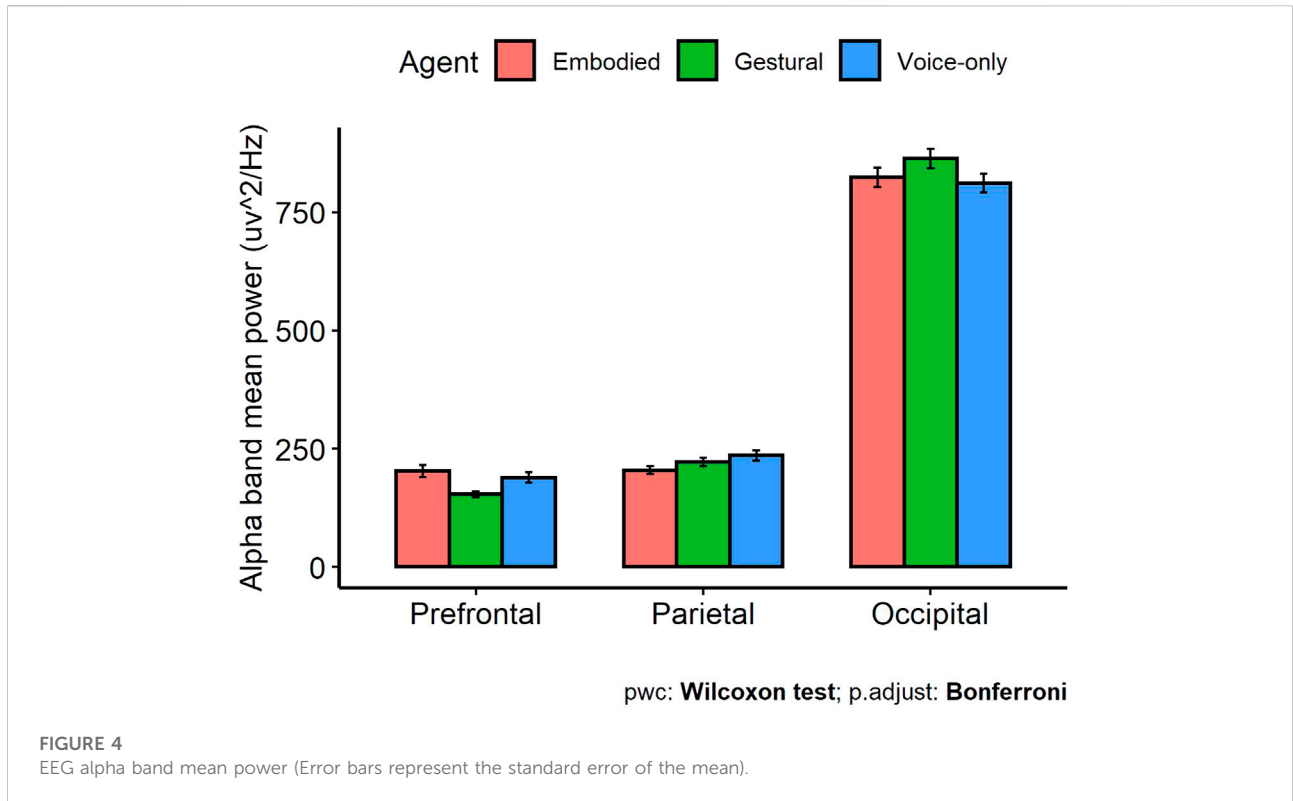


FIGURE 4
EEG alpha band mean power (Error bars represent the standard error of the mean).

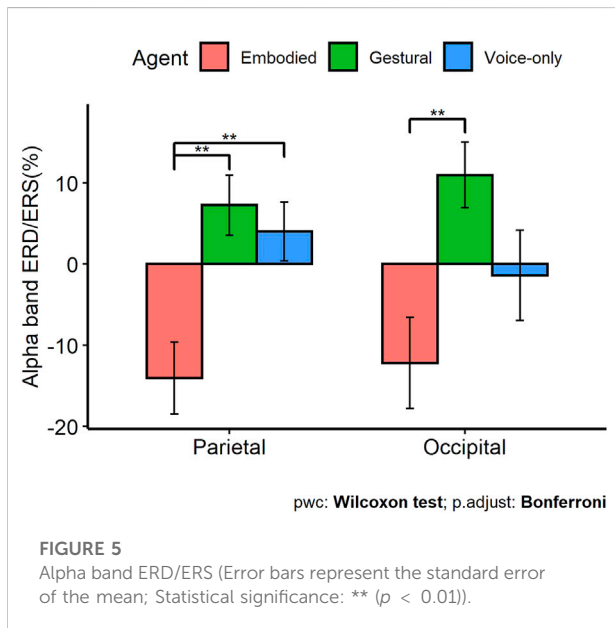


FIGURE 5
Alpha band ERD/ERS (Error bars represent the standard error of the mean; Statistical significance: ** ($p < 0.01$)).

A one-way repeated measure ANOVA was performed on the weighted NASA TLX score. We found a significant difference ($F(2, 20) = 7.79, p = 0.003, \eta^2 = 0.08$) among the three conditions. Post-hoc test showed the workload was significantly lower in the conditions of

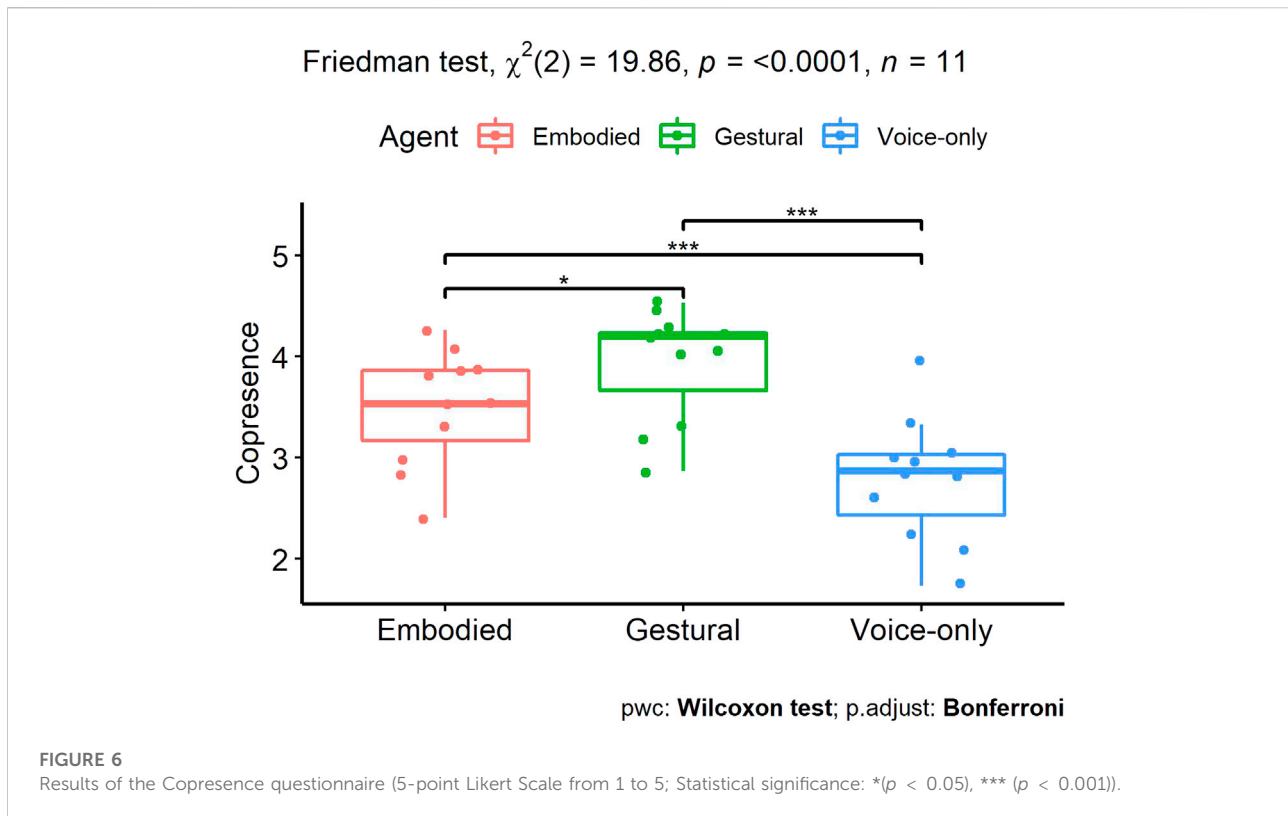
Embodied Agent ($M = 56.87, SD = 15.89, p = 0.008$) and Gestural Agent ($M = 58.02, SD = 12.89, p = 0.048$) compared with the Voice-only ($M = 65.95, SD = 12.94$) Agent. No significant difference was found between the Gestural and Embodied Agent (see Figure 7).

The Friedman test followed by Kendall’s W test on the SMEQ data on the three agent conditions showed significant difference ($\chi^2(2) = 12, p < 0.01$) with large effect size ($W = 0.55$). The post-hoc analysis with Wilcoxon tests revealed both the SMEQ results for the Embodied condition ($Z = -2.39, p = 0.017$) and Gestural ($Z = -2.33, p = 0.02$) condition had significant lower rating scores than the Voice-only condition (see Figure 8).

9.4 Discussion

9.4.1 Task performance, copresence and workload

Our findings on task performance, copresence and workload partially support H1. Visually embodied agents could create a stronger copresence feeling and achieve a lower task load without influencing the task performance, compared with the Voice-only Agent (baseline). Furthermore, the Gestural Agent received a higher rating score on copresence than the Embodied Agent and can reduce task load in human-agent collaborative decision-making tasks.



We could not find any significant difference in the task performance. This could be because the small sample size and potential learning effects. Several participants reported they could guess and remember a few general sequences the virtual agent suggested repeatedly in previous trials. Interestingly, the descriptive statistics results of task performance were identical to the results of copresence.

Participants had the strongest feeling of copresence with the Gestural Agent, followed by the Embodied Agent, compared with the baseline. This could be because the non-verbal cues like gaze, head movement, and gestures could improve the IVAs' naturalness and believability (Hanna and Richards, 2016). Most participants reported that they liked the Gestural Agent best in the post-experiment interview. One participant said, "the gestural agent's pointing gestures coupled with its voice content makes it more natural and real". Another reported, "the pointing gesture synchronized with the voice makes that agent feel relatively more vivid".

Both NASA-TLX and SMEQ revealed that the workload in the Gestural and Embodied agent conditions was significantly lower than that in the Voice-only agent condition. This agrees with prior work (Kim et al., 2020), where an IVA in AR with both gestures and body movement reduced the task load to some extent. The task load in our experiment mainly came from determining the order of 15 items and understanding the suggestions from the virtual agents. All three agents provided correct suggestions for participants, while the gestural agent

looked at the suggested item and pointed it while speaking, which may reduce task load caused by understanding the agent's speech contents. One participant commented that "the pointing gesture can make me understand easily which item he (gestural agent) is talking about and quickly locate that item".

9.4.2 Alpha band mean power and cognitive load

Our findings on alpha band mean power fails to support H2. The cognitive load of our research task mainly originated from making decisions on determining orders of the 15 desert survival items (intrinsic cognitive load) and interacting with the virtual agent (extraneous cognitive load). We will analyze these two types of cognitive load in more detail and discuss possible reasons why the alpha band power failed to show significant difference.

The cognitive load of making decisions in our research task could be caused by retrieving experiences and knowledge of using those items, comparing their importance and planning their orders. Prefrontal cortex was showed to be related to LTM retrieval (Buckner and Petersen, 1996) and relational reasoning (Waltz et al., 1999). Therefore, we expected to capture this type of cognitive load using the Fp1 and Fp2 EEG sites. However, the alpha band power obtained from these sites showed significant difference among the three agent conditions. This could be because the alpha band mean power was calculated with signals cropped by the time window when the virtual agent was giving suggestions

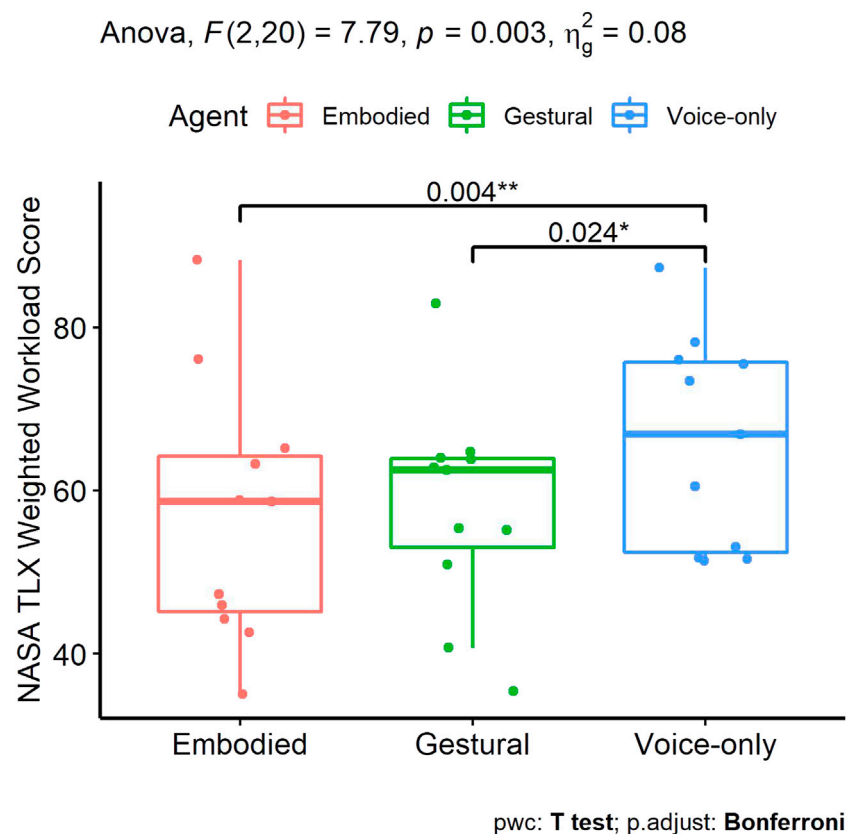


FIGURE 7

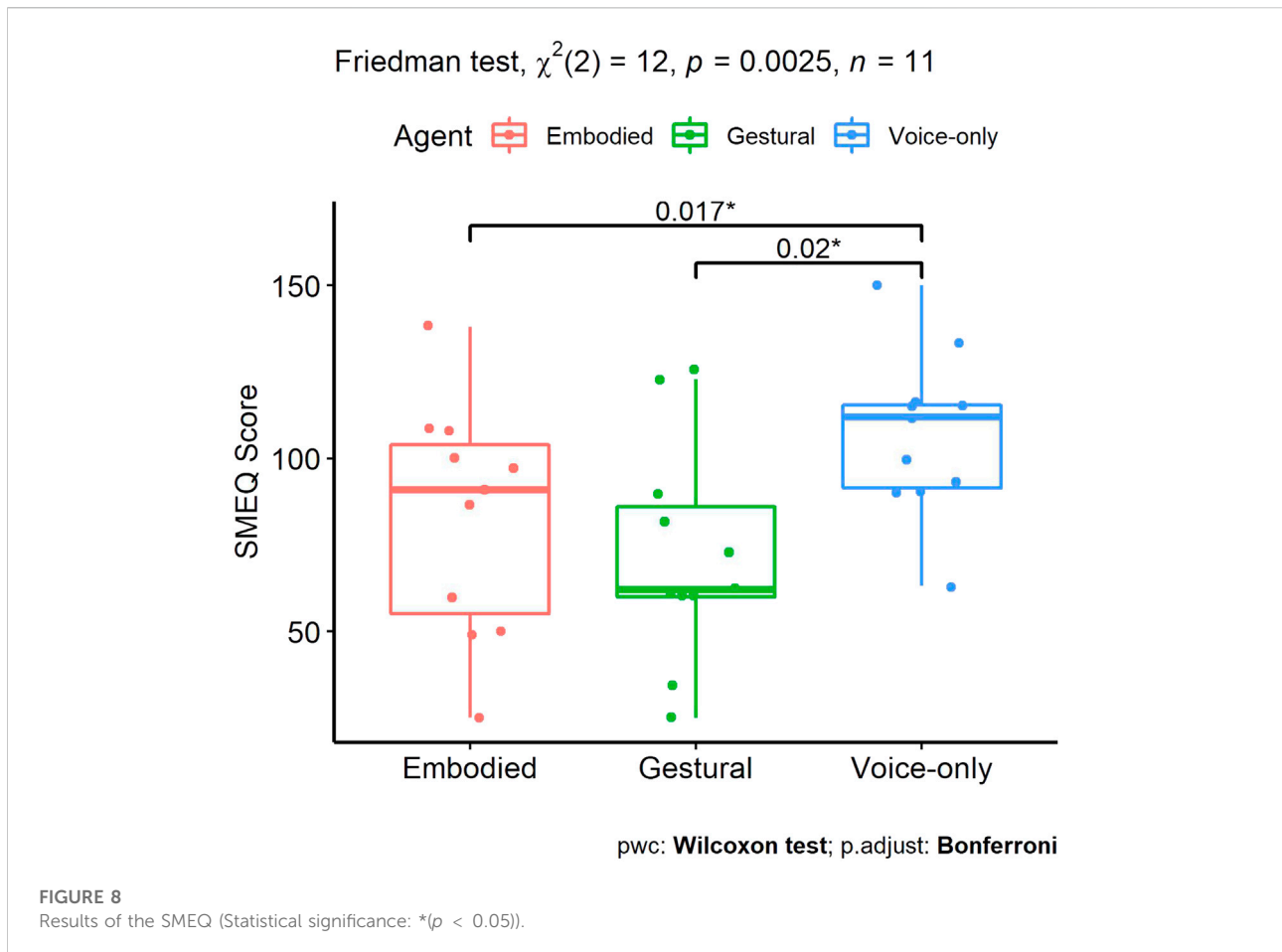
Results of the NASA TLX weighted workload (Statistical significance: $*p < 0.05$, $**p < 0.01$).

whereas participants might also mentally compare the items' importance while adjusting their orders.

The cognitive load originated from interacting with virtual agents mainly comprises processing the coupled visual-audio stimuli emitted by the virtual agent and visually searching the object that the agent were talking about. Since parietal lobe was proved to be related to selective attention (Behrmann et al., 2004) and occipital lobe was thought to be in charge of visual processing (Wandell et al., 2007), we expected to reflect this type of cognitive load with the signals captured from P3, P4, and Pz EEG sites located at parietal lobe and the O1 and O2 sites located at occipital lobe. However, neither the parietal nor the occipital alpha band mean power showed significant difference among the three agent conditions. This could be because both the Embodied and the Gestural agent only kept tiny and monotonous body movement animation while talking. People could roughly locate the suggested item with the help of gazing direction when the Embodied agent and Gestural agent started speaking. They would precisely lock the suggested item when the agent's speech reveals more information about that item. However, it was also not hard to tell which object the virtual agent was talking about with

speech only. Once they confirmed the location of the suggested object, the only thing they expected from the virtual agent would be the direction of adjusting. Therefore, people might mainly focus on the speech content while not the agent body after quickly locating the suggested item, which could explain why there was no significant difference in both parietal and occipital lobe among the three conditions.

Although the alpha band power, NASA TLX and SMEQ were all used to measure cognitive load, it is interesting to note the difference between our alpha band power results and the two subjective cognitive measures. First, the alpha power was calculated with signals when the virtual agent was speaking whereas the questionnaires were filled after each trial. As we have mentioned, the alpha band power results might miss some decision-making cognitive load when the agent was not talking. However, the subjective measures were about the general feeling of work load while performing the task. Second, the different impressions on the three types of agents might influence their feelings of the overall workload which could not be revealed by our EEG alpha band power results. For example, as we have mentioned in Section 9.4.1 people might like the Embodied and Gestural agent because of their visual representations and the



indicating cues (gaze in both condition and the pointing gesture in Gestural condition). This subjective feelings on visually embodied agents might influence the subjective assessment on overall cognitive load. This analysis could be supported by the results of alpha band ERD/ERS discussed in the following section.

9.4.3 Alpha band the event related (De-)synchronization

Our alpha band ERD/ERS results partially support the H2. For the ERD/ERS in alpha band, we found significant difference in both parietal and occipital lobe. When the agent started speaking, the alpha band power increased (ERS) in the Embodied Agent condition, whereas that decreased (ERD) in both Gestural and Voice-only Agents condition. In other words, the Gestural agent and Voice-only Agents required more attentional demands compared to the Embodied Agent. Before each agent started speaking, participants would wait for 2 ~3 s without any sudden audio or visual changes. Therefore, we assumed there were more alpha oscillations in participants' brain activity before the agent started speaking.

When the Embodied Agent started speaking, the speech was matched with the virtual body and thus being expected by the

participants. However, the Voice-only Agent had no virtual body and the sudden speech stimulus might cause the alpha suppression. One participant commented that “the Voice-only Agent’s sudden speech makes it feels abrupt”.

When the Gestural Agent started speaking, the coupled pointing gesture attracted more visual attention and thus caused the alpha suppression which could explain the ERD in the Gestural Agent condition. Although the pointing gesture increased the copresence of the Gestural Agent compared to the Embodied Agent (see results for Copresence), it could cause more attention when the Gestural Agent started speaking.

9.4.4 Limitations

Our research had some limitations. First, we had difficulty finding female participants due to their reluctance to wash their hair after having EEG gel placed on it. We also recruited participants from a mechanical engineering school where men were overwhelmingly more available than women. Secondly, similar to Gupta et al. (2019), we also encountered difficulties mounting the HTC VIVE on the EEG cap. Fastening the VR headset squeezes prefrontal, parietal, and occipital skull areas and

may cause some electrodes to slide or even lose contact with the scalp. This may result in poor EEG signals.

10 Conclusion and future work

In this paper, we explored the effects of virtual agent's interaction methods on users' brain activity. We reported a within-subject pilot study where participants were asked to work collaboratively with a virtual agent on a decision-making task in VR environment. Three types of virtual agents were designed by varying the representation: 1) a Voice-only Agent communicating through speech only, 2) an Embodied Agent with a virtual body interacting through both speech and gaze, and 3) a Gestural Agent embodied with a virtual body allowing to express information through speech, gaze and a pointing gesture. We recorded users' EEG signals with which we calculated the alpha band power and ERD/ERS, which were proved to be effective in reflecting cognitive load. Except for the EEG signals, we also collected subjective ratings of cognitive load and feelings of copresence, and task performance to reveal the possible significant difference among the three agent conditions from different perspectives. No significant difference was found in the task performance. Subjective results indicate that imbuing an agent with a visual body in VR and task-related gestures could reduce task load and improve copresence in human-agent collaborative decision making. The alpha band power showed no significant difference among the three agent conditions in prefrontal cortex, parietal lobe and occipital lobe. The result of alpha ERD/ERS in parietal and occipital lobe revealed that the coherence in agent's speech and body representation was important. And the non-verbal cues like the task related pointing gesture required more attentional demand, even though it was helpful in locating objects and improving overall copresence.

Overall, virtual agent designers should be careful in designing non-verbal cues for virtual agents by making trade-offs between the benefits of such cues and the possible distractions or increased attentional demands. We argue that interacting with virtual agent can cause complex changes in users' brain activity. Understanding the impact of interacting with virtual agents in AR and VR on users' physiological states like brain activity could help to improve the design of such agents. We call for more attention to be paid on this important topic.

In the future, we plan to investigate if there is any difference in other EEG bands using VR headsets with better EEG fit. Furthermore, We will look into IVAs' adaptive behaviors according to users' physiological states captured by various physiological sensors like EEG, galvanic skin response, and heart-rate variability simultaneously. Last, comparing the effects of interacting with a virtual agent and an actual human on users' physiological states is also an interesting topic.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Cyber Physical Interaction Lab and 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.

Author contributions

ZC contributed to the conception, design, development, conduct and report of the analysis of the study. HB and MB supervised the study as well as part of the report of the study. ZC, LZ, and WH contributed to the conduct of the study. KG contributed to the experiment design and EEG data analysis. ZC wrote the draft of the manuscript. HB, LZ, and MB contributed to the manuscript revision and read.

Funding

This paper is sponsored by the China Scholarship Council (CSC NO. 202008320270).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Antonenko, P., Paas, F., Grabner, R., and Van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educ. Psychol. Rev.* 22, 425–438. doi:10.1007/s10648-010-9130-y
- Behrmann, M., Geng, J. J., and Shomstein, S. (2004). Parietal cortex and attention. *Curr. Opin. Neurobiol.* 14, 212–217. doi:10.1016/j.conb.2004.03.012
- Buckner, R. L., and Petersen, S. E. (1996). What does neuroimaging tell us about the role of prefrontal cortex in memory retrieval? *Semin. Neurosci.* 8, 47–55. Academic Press. doi:10.1006/smns.1996.0007
- de Melo, C. M., Kim, K., Norouzi, N., Bruder, G., and Welch, G. (2020). Reducing cognitive load and improving warfighter problem solving with intelligent virtual assistants. *Front. Psychol.* 11, 554706. doi:10.3389/fpsyg.2020.554706
- Dey, A., Chatburn, A., and Billinghurst, M. (2019). “Exploration of an eeg-based cognitively adaptive training system in virtual reality,” in 2019 IEEE conference on virtual reality and 3d user interfaces (vr), Osaka, Japan, March 23–27, 2019 (IEEE), 220–226.
- Fink, A., and Benedek, M. (2014). Eeg alpha power and creative ideation. *Neurosci. Biobehav. Rev.* 44, 111–123. doi:10.1016/j.neubiorev.2012.12.002
- Fink, A., Rominger, C., Benedek, M., Perchtold, C. M., Papousek, I., Weiss, E. M., et al. (2018). Eeg alpha activity during imagining creative moves in soccer decision-making situations. *Neuropsychologia* 114, 118–124. doi:10.1016/j.neuropsychologia.2018.04.025
- Freeman, W. J. (2002). “Making sense of brain waves: the most baffling frontier in neuroscience,” in *Biocomputing* (Springer), 1–23.
- Gerry, L., Ens, B., Drogemuller, A., Thomas, B., and Billinghurst, M. (2018). “Levity: A virtual reality system that responds to cognitive load,” in Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, Montréal, Canada, April 21–26, 2018, 1–6.
- Gevins, A., and Smith, M. E. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. *Theor. Issues Ergon. Sci.* 4, 113–131. doi:10.1080/14639220210159717
- Gupta, K., Hajika, R., Pai, Y. S., Duenser, A., Lochner, M., and Billinghurst, M. (2019). “In ai we trust: Investigating the relationship between biosignals, trust and cognitive load in vr,” in 25th ACM Symposium on Virtual Reality Software and Technology, Parramatta, NSW, Australia, November 12–15, 2019, 1–10.
- Gupta, K., Hajika, R., Pai, Y. S., Duenser, A., Lochner, M., and Billinghurst, M. (2020). “Measuring human trust in a virtual assistant using physiological sensing in virtual reality,” in 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Atlanta, Georgia, USA, March 22–26, 2020 (IEEE), 756–765.
- Haesler, S., Kim, K., Bruder, G., and Welch, G. (2018). “Seeing is believing: improving the perceived trust in visually embodied alexa in augmented reality,” in 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Munich, Germany, March 16–20, 2018 (IEEE), 204–205.
- Hanna, N. H., and Richards, D. (2016). *Human-agent teamwork in collaborative virtual environments*. Ph.D. thesis. Sydney, Australia: Macquarie University. Available at: <http://hdl.handle.net/1959.14/1074981>.
- Hantono, B. S., Nugroho, L. E., and Santosa, P. I. (2016). “Review of augmented reality agent in education,” in 2016 6th International Annual Engineering Seminar (InAES), Yogyakarta, Indonesia, August 1–3, 2016 (IEEE), 150–153.
- Harmony, T., Fernández, T., Silva, J., Bernal, J., Díaz-Comas, L., Reyes, A., et al. (1996). Eeg delta activity: an indicator of attention to internal processing during performance of mental tasks. *Int. J. Psychophysiol.* 24, 161–171. doi:10.1016/s0167-8760(96)00053-0
- Hart, S. G., and Staveland, L. E. (1988). “Development of nasa-tlx (task load index): Results of empirical and theoretical research,” in *Advances in psychology* (Elsevier), 52, 139–183.
- Hollender, N., Hofmann, C., Deneke, M., and Schmitz, B. (2010). Integrating cognitive load theory and concepts of human-computer interaction. *Comput. Hum. Behav.* 26, 1278–1288. doi:10.1016/j.chb.2010.05.031
- Holm, A., Lukander, K., Korpela, J., Sallinen, M., and Müller, K. M. (2009). Estimating brain load from the eeg. *Sci. World J.* 9, 639–651. doi:10.1100/tsw.2009.83
- Holz, T., Campbell, A. G., O’Hare, G. M., Stafford, J. W., Martin, A., and Dragone, M. (2011). Mira—mixed reality agents. *Int. J. Human-comput. Stud.* 69, 251–268. doi:10.1016/j.ijhcs.2010.10.001
- Jing, A., May, K., Lee, G., and Billinghurst, M. (2021). Eye see what you see: Exploring how bi-directional augmented reality gaze visualisation influences co-located symmetric collaboration. *Front. Virtual Real.* 2, 79. doi:10.3389/frvir.2021.697367
- Kevin, S., Pai, Y. S., and 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, Tokyo, Japan, November 28–December 1, 2018, 1–2.
- Kim, K., Boelling, L., Haesler, S., Bailenson, J., Bruder, G., and Welch, G. F. (2018). “Does a digital assistant need a body? the influence of visual embodiment and social behavior on the perception of intelligent virtual agents in ar,” in 2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Munich, Germany, October 16–20, 2018 (IEEE), 105–114.
- Kim, K., de Melo, C. M., Norouzi, N., Bruder, G., and Welch, G. F. (2020). “Reducing task load with an embodied intelligent virtual assistant for improved performance in collaborative decision making,” in 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Atlanta, Georgia, America, March 22–26, 2020 (IEEE), 529–538.
- Klimesch, W., Pfurtscheller, G., and Schimke, H. (1992). Pre-and post-stimulus processes in category judgement tasks as measured by event-related desynchronization (erd). *J. Psychophysiol.* 6 (3), 185–203.
- Klimesch, W. (1997). Eeg-alpha rhythms and memory processes. *Int. J. Psychophysiol.* 26, 319–340. doi:10.1016/s0167-8760(97)00773-3
- Klimesch, W. (1999). Eeg alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Rev.* 29, 169–195. doi:10.1016/s0165-0173(98)00056-3
- Kumar, N., and Kumar, J. (2016). Measurement of cognitive load in hci systems using eeg power spectrum: an experimental study. *Procedia Comput. Sci.* 84, 70–78. doi:10.1016/j.procs.2016.04.068
- Lang, W., Lang, M., Kornhuber, A., Diekmann, V., and Kornhuber, H. (1988). Event-related eeg-spectra in a concept formation task. *Hum. Neurobiol.* 6, 295–301.
- Lécuyer, A., Lotte, F., Reilly, R. B., Leeb, R., Hirose, M., and Slater, M. (2008). Brain-computer interfaces, virtual reality, and videogames. *Computer* 41, 66–72. doi:10.1109/mc.2008.410
- Li, C., Androulakaki, T., Gao, A. Y., Yang, F., Saikia, H., Peters, C., et al. (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, Sydney, NSW, Australia, November 5–8, 2018, 191–196.
- Luck, S. J. (2014). *An introduction to the event-related potential technique*. Chicago: MIT press.
- Malach, R., Reppas, J., Benson, R., Kwong, K., Jiang, H., Kennedy, W., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc. Natl. Acad. Sci. U. S. A.* 92, 8135–8139. doi:10.1073/pnas.92.18.8135
- Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi:10.1146/annurev.neuro.24.1.167
- Miller, M. R., Jun, H., Herrera, F., Yu Villa, J., Welch, G., and Bailenson, J. N. (2019). Social interaction in augmented reality. *PLoS one* 14, e0216290. doi:10.1371/journal.pone.0216290
- Mustafa, M., Guthe, S., Tauscher, J.-P., Goesele, M., and Magnor, M. (2017). “How human am i? eeg-based evaluation of virtual characters,” in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, Colorado, USA, May 6–11, 2017, 5098–5108.
- Norouzi, N., Kim, K., Hochreiter, J., Lee, M., Daher, S., Bruder, G., et al. (2018). “A systematic survey of 15 years of user studies published in the intelligent virtual agents conference,” in Proceedings of the 18th international conference on intelligent virtual agents, Sydney, NSW, Australia, November 5–8, 2018, 17–22.
- Norouzi, N., Bruder, G., Belna, B., Mutter, S., Turgut, D., and Welch, G. (2019). “A systematic review of the convergence of augmented reality, intelligent virtual agents, and the internet of things,” in *Artificial intelligence in IoT*, 1–24.
- Norouzi, N., Kim, K., Bruder, G., Erickson, A., Choudhary, Z., Li, Y., et al. (2020). “A systematic literature review of embodied augmented reality agents in head-mounted display environments,” in International Conference on Artificial Reality and Telexistence-Eurographics Symposium on Virtual Environments, Orlando, Florida, December 2–4, 2020, 101–111.
- Oostenveld, R., and Praamstra, P. (2001). The five percent electrode system for high-resolution eeg and erp measurements. *Clin. Neurophysiol.* 112, 713–719. doi:10.1016/s1388-2457(00)00527-7
- Paas, F., Renkl, A., and Sweller, J. (2003a). Cognitive load theory and instructional design: Recent developments. *Educ. Psychol.* 38, 1–4. doi:10.1207/s15326985ep3801_1
- Paas, F., Tuovinen, J. E., Tabbers, H., and Van Gerven, P. W. (2003b). “Cognitive load measurement as a means to advance cognitive load theory,” in *Educational psychologist* (Routledge), 63–71.
- Paas, F., Renkl, A., and Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instr. Sci.* 32, 1–8. doi:10.1023/b:truc.0000021806.17516.d0

- Pimentel, D., and Vinkers, C. (2021). Copresence with virtual humans in mixed reality: The impact of contextual responsiveness on social perceptions. *Front. Robot. AI* 8, 634520. doi:10.3389/frrobot.2021.634520
- Ramchurn, S. D., Wu, F., Jiang, W., Fischer, J. E., Reece, S., Roberts, S., et al. (2016). Human-agent collaboration for disaster response. *Auton. Agent. Multi. Agent. Syst.* 30, 82–111. doi:10.1007/s10458-015-9286-4
- Reinhardt, J., Hillen, L., and Wolf, K. (2020). “Embedding conversational agents into ar: Invisible or with a realistic human body?” in Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction, Sydney NSW Australia, February 9–12, 2020, 299–310.
- Sauro, J., and Dumas, J. S. (2009). “Comparison of three one-question, post-task usability questionnaires,” in Proceedings of the SIGCHI conference on human factors in computing systems, Boston MA, USA, April 4–9, 2009, 1599–1608.
- Stipacek, A., Grabner, R., Neuper, C., Fink, A., and Neubauer, A. (2003). Sensitivity of human eeg alpha band desynchronization to different working memory components and increasing levels of memory load. *Neurosci. Lett.* 353, 193–196. doi:10.1016/j.neulet.2003.09.044
- Suzuki, Y., Galli, L., Ikeda, A., Itakura, S., and Kitazaki, M. (2015). Measuring empathy for human and robot hand pain using electroencephalography. *Sci. Rep.* 5, 15924–15929. doi:10.1038/srep15924
- Sweller, J., Van Merriënboer, J. J., and Paas, F. G. (1998). Cognitive architecture and instructional design. *Educ. Psychol. Rev.* 10, 251–296. doi:10.1023/a:1022193728205
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Sci.* 12, 257–285. doi:10.1207/s15516709cog1202_4
- Teplan, M. (2002). Fundamentals of eeg measurement. *Meas. Sci. Rev.* 2, 1–11.
- Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santos, M., et al. (1999). A system for relational reasoning in human prefrontal cortex. *Psychol. Sci.* 10, 119–125. doi:10.1111/1467-9280.00118
- Wandell, B. A., Dumoulin, S. O., and Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron* 56, 366–383. doi:10.1016/j.neuron.2007.10.012
- Wang, I., and Ruiz, J. (2021). Examining the use of nonverbal communication in virtual agents. *Int. J. Human-Computer. Interact.* 37, 1648–1673. doi:10.1080/10447318.2021.1898851
- Wang, I., Smith, J., and Ruiz, J. (2019). “Exploring virtual agents for augmented reality,” in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, May 4–9, 2019, 1–12.
- Ye, Z.-M., Chen, J.-L., Wang, M., and Yang, Y.-L. (2021). “Paval: Position-aware virtual agent locomotion for assisted virtual reality navigation,” in 2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Bari, Italy, October 4–8, 2021 (IEEE), 239–247.
- Zhang, L., Wade, J., Bian, D., Fan, J., Swanson, A., Weitlauf, A., et al. (2017). Cognitive load measurement in a virtual reality-based driving system for autism intervention. *IEEE Trans. Affect. Comput.* 8, 176–189. doi:10.1109/taffc.2016.2582490
- Zijlstra, F., and Van Doorn, L. (1985). *The construction of a scale to measure perceived effort*. Delft, NL: Delft University of Technology.