Cognitive Impact of Social Virtual Reality: Audience and Mere-Presence Effect of Virtual Companions

Olga Sutskova [1, 2], Atsushi Senju [1], Tim J Smith [2,3]

[1] Research Centre For Child Mental Development, Hamamatsu University School Of Medicine, 431-3125, Japan.

[2] Centre of Brain and Cognitive Development, Birkbeck University of London, WC1E 7HX, United Kingdom.

[3] Creative Computing Institute, University of the Arts London, London WC1V 7EY, United Kingdom.

Cognitive Impact of Social Virtual Reality: Audience and Mere-Presence Effect of Virtual Companions

Olga Sutskova [1,2], Atsushi Senju [1], Tim J Smith [2,3]

[1] Research Centre For Child Mental Development, Hamamatsu University School Of Medicine, 431-3125, Japan.

[2] Centre of Brain and Cognitive Development, Birkbeck University of London, WC1E 7HX, United Kingdom.

[3] Creative Computing Institute, University of the Arts London, London WC1V 7EY, United Kingdom.

Abstract

Sharing experiences with others is an important part of everyday life. Immersive virtual reality (IVR) promises to simulate these experiences. However, whether IVR elicits a similar level of social presence as measured in the real world is unclear. It is also uncertain whether AI-driven virtual humans (agents) can elicit a similar level of meaningful social co-presence as people-driven virtual-humans (avatars). The current study demonstrates that both virtual human types can elicit a cognitive impact on a social partner. The current experiment tested participants' cognitive performance changes in the presence of virtual social partners by measuring the social facilitation effect (SFE). The SFE-related performance change can occur through either vigilance-based mechanisms related to other people's co-presence (known as the mere presence effect: MPE) or reputation management mechanisms related to other people's monitoring (the audience effect: AE). In this study, we hypothesised AE and MPE as distinct mechanisms of eliciting SFE. Firstly, we predicted that, if head-mounted IVR can simulate sufficient co-presence, any social companion's visual presence would elicit SFE through MPE. The results demonstrated that companion presence decreased participants' performance irrespective of whether AI or human-driven. Secondly, we predicted that monitoring by a human-driven, but not an AI-driven, companion would elicit SFE through AE. The results demonstrated that monitoring by a human-driven companion affected participant performance more than AI-driven, worsening performance marginally in accuracy and significantly in reaction times. We discuss how the current results explain the findings in prior SFE in virtual-world literature, and map out future considerations for social-IVR testing, such as participants virtual self-presence, and affordances of physical and IVR testing environment.

Cognitive Impact of Social Virtual Reality: Audience and Mere-Presence Effect of Virtual Companions

Introduction

Humans are affected by other people in their shared environment (Boothby et al., 2017). The lengthy restrictions on in-person interaction during the years of the COVID-19 pandemic have emphasised the importance of in-person experience sharing. A growing requirement for a genuine sense of social co-presence with others urged the entertainment and social media industries to seek solutions for simulating these experiences remotely, re-igniting the curiosity in immersive virtual reality (IVR) technologies.

The significance of the IVR experience lies in its ability to simulate a sense of agency and self-presence in an interactive virtual world that surrounds its user with a 360-degree digital environment. The IVR experience often perceptually removes the participant from their real-world surroundings. Through immersive presence, the participants can act upon the virtual environment similarly to the real world and exceed the rules of physical reality. Since IVR became commercially available, there has been a growing interest in how immersive reality affordances can be used in psychological research (Zender et al., 2019). The findings are currently promising. For example, prior research has shown that embodied interaction with the immersive physical environment can positively influence cognitive performance and foster more creative problem-solving (Wang et al., 2018), facilitating intuitive cognition (Eskinazi & Giannopulu, 2021). In agreement with an original proposal by Blascovich and colleagues (2002), researchers are now finding that testing in IVR can indeed offer a more ecologically valid solution to testing classic paradigms, which until recently were limited to artificial screen-based stimuli (Olk et al., 2018).

The effects explored in IVR experimentally are often non-social. However, with growing interest in immersive social interaction, it is now crucial to establish and test the baseline mechanism of immersive virtual social reality. This is the aim of the current paper.

The social IVR encloses its users within a computer-generated interactive world, developed around the user and their virtual companions. In contrast to desktop-based social interaction, during which the user is located within their home environment viewing the companion on-screen, in IVR the user virtually co-inhabits the immersive environment alongside their social companion. Within the immersive space, the companion and participants can directly approach and virtually engage with one another's virtual representation in the simulated environment, leading to higher levels of co-presence in the same environment than desktop-based communication (Makransky et al., 2019; Oh et al., 2018; Witmer & Singer, 1998). However, whether this co-immersion in IVR elicits a meaningful sense of co-presence, equivalent to real-world interaction, is yet to be determined.

The current study tests whether a classic social phenomenon, reported during a real-world interaction, can be also elicited within the IVR. In particular, the current study focuses on the social facilitation effect (SFE), which manifests as the change in individuals' performance when they perform within a social context, in contrast to when performing alone (Bond & Titus, 1983; Guerin, 1986; Rajecki, 2010; Zajonc, 1965). As per canonical SFE observed in real-world social contexts, we expect the canonical SFE within IVR to follow a similar pattern of influence on performance – the cognitive performance will improve (facilitation) when they perform an easy task and decrease (inhibition) when they perform more difficult tasks, in presence of an immersive virtual companion in contrast to when performing alone. Given the importance of cognitive performance during remote interaction in education and work settings, including IVR, it is important to understand whether cognitive performance changes driven by SFE's are also observed due to the presence of social agents within virtual scenarios. However, previous studies attempting to measure cognitive SFE in an immersive

setting have resulted in mixed findings, with most studies not showing a canonical pattern of the SFE, but only showing performance inhibition without facilitation (Emmerich & Masuch, 2016; Hoyt et al., 2003; Zanbaka et al., 2007; Terrier et al., 2019) and one resulting in null findings (Hayes et al., 2010).

The most straightforward explanation of the discrepancy in prior virtual SFE findings could be that different studies used different types of companions and that some types of virtual companions accompanying the participant might not be as socially influential as others. It is often assumed that the human-driven companion's (i.e., an avatar's) ability to mentalise from another person's perspective, renders it more socially meaningful to the participant, in contrast to an AI-driven companion (i.e., agent; Blascovich, 2002), which would not possess a similar capacity to mentalise. However, the findings on whether the SFE within an immersive context is affected by the companion's agency (human or AI-driven) are also inconclusive. Some reports demonstrate performance inhibition in the presence of avatars but not of agents (Hoyt et al., 2003), whilst others report performance inhibition for both (Emmerich & Masuch, 2016; Zanbaka et al., 2007).

If the social influence on performance cannot be explained just through the companions' agency (the mind behind the companion), it is possible that the performance change could also be driven by the companion's interactive and visual factors. Prior work suggests that a virtual companion's human likeness, i.e., whether the companion looks and acts more humanlike versus non-humanoid, is a significant predictor of companions being treated as other people (Nass et al., 1994; Nass & Moon, 2000; Von Der Pütten et al., 2010). However, the SFE, which manifested as inhibition without facilitation, has also been reported for robotlike virtual companions (Emmerich & Masuch, 2016). This finding could be explained by emerging theoretical work in human-agent interaction that proposes the significance of multimodality in interaction, emphasising the importance of virtual communication context rather than just the companion visual attributes. In their virtual interaction model, Klowait and Erofeeva (2021) propose that when the participant learns of the contingencies within the social virtual interaction context, the agent companions' sufficient humanlike interactive properties can overwrite the insufficient visual humanlike attributes. Therefore, even if the agent companion is not sufficiently human-like in appearance but represents a sufficient human-like function to the participant, it will be interacted with similarly to a real person due to pragmatic heuristics.

A systematic experimental contrast of humanoid versus non-humanoid presence has not been tested in the framework of immersive virtual SFE. Therefore, it is still not clear whether the agency of the virtual companion, its physical attributes (visual human-likeness, human-like function), or the combination of these two, influences a meaningful sense of immersive co-presence and elicits the SFE. And if virtual SFE is elicited, which cognitive mechanisms are responsible for these effects.

We propose that investigating SFE through its hypothesised mechanisms can provide an original perspective on how different companion types might elicit SFE within the IVR and elucidate the driving forces behind the SFE itself. To the best of the authors' knowledge, the prior published studies investigating SFE within a virtual context, measured the SFE as a general effect, without considering the SFE's underlying mechanisms. Considering that virtual companionship already transcends entertainment industries, heading towards well-being, work, and education (Gasteiger et al., 2021; Huang & Rust, 2018; Iqbal et al., 2021), it is important to investigate which virtual companions are truly socially meaningful, and what are the consequences of interactions with them.

The real-world (not virtual) SFE has been theorised as being due to two potential cognitive mechanisms (Guerin & Innes, 1984), either our awareness of the mere presence of another person in the environment (the mere presence effect: MPE), or our sense of being observed by another person (the audience effect: AE). In the SFE literature, the MPE and AE are often researched separately and are therefore referred to as "effects" on their own, even if both result in SFE under different social

conditions. To resolve the confusion, in the current paper, the terms of MPE and AE are being used as distinct mechanisms which elicit SFE as the outcome.

It is important to understand the distinction between the MPE and AE as cognitive mechanisms because there are fundamental differences in cognitive processes related to how social contexts elicit the SFE. The MPE is based on the registration of physical uncertainty and vigilance over another person's presence in the same environment. Critically, the MPE is agnostic of whether the other person is explicitly monitoring/observing the participant (Green & Bushman, 1989). In contrast, AE is driven by the belief of being watched and potentially judged whilst performing a task (Hamilton & Lind, 2016), for example through a camera, irrespective of whether the observer is visually co-present in the shared environment (Dumontheil et al., 2016; Wolf et al., 2015).

Considering the differences between the MPE and the AE, it is hypothesised that through the MPE, the companion's physical embodied (bodily) presence alone can elicit the SFE, by arousing participants' vigilance state, irrespective of the companion mind property (agency: human- or AI-driven). Whilst through the AE, being watched by a mentalising agent (such as a human-controlled avatar in contrast to an AI-controlled agent) is hypothesised to be sufficient to elicit the SFE, irrespective of the companion's visual body presence in the same environment. In the real-world interaction scenario, it is challenging to separate these two properties of a companion, especially the co-presence of an interactive partner who lacks the ability to mentalise. However, IVR provides a unique opportunity to experimentally dissociate these two properties, isolating the companion's virtual body from their mind.

To disentangle how virtual companions are perceived in IVR, the consideration for the categorical differences between the two cognitive mechanisms underlying SFE (MPE and AE) could be crucial, revealing participants' underlying social cognition when exposed to different types of virtual companions. Up until now, the IVR studies used SFE as an umbrella term, merging its two mechanisms, MPE and AE, and observing SFE as an outcome. These experimental choices potentially enabled the participants' subjective beliefs about companions' mentalising properties and agency behind their actions to vary freely. If this is indeed the case, it is still not clear whether the inhibitory effect reported in prior immersive SFE literature is driven by beliefs of companion agency (related to the AE) or their visual human-likeness during immersive co-presence (not related to the AE but possibly related to the MPE, as we discuss below).

Based on the MPE, it is hypothesised that if the companion's embodied co-presence in an immersive space is socially meaningful irrespective of their agency, the social virtual companion co-presence should be sufficient to elicit SFE. Indeed, a review of virtual interaction (Oh et al., 2018) shows that communicative social AI agents can be engaging even if they are not in visually humanoid form. However, more visually humanlike companions are believed to be most socially impactful (Appel et al., 2012; Nass et al., 1994; Von Der Pütten et al., 2010). Interestingly, based on the MPE, it is the sense of social embodied presence of a companion in the shared space that drives the effect, irrespective of whether the companion is attentive to the participant (Rajecki et al., 1977). Therefore, whether the companion has a human mind and is able to appraise the human partner is irrelevant for the MPE to elicit the SFE. If so, both the avatar and agent co-presence should be sufficiently and equally impactful, as per MPE. Whether the companion's humanoid form, irrespective of agency, is important for virtual MPE to emerge, is directly tested in the current study.

In contrast to the MPE, the AE is hypothesised to depend on the companion having the mental capacity to monitor the participant. If a participant's anticipation of social judgement indeed drives the effect (Guerin & Innes, 1984), an avatar's mentalising capacity renders it more socially meaningful than an agent, which is without mentalising capacity when the participant believes their performance is monitored. Indeed, the threshold model of social influence (TMSI: Blascovich, 2002) predicts that when evaluation is central to interaction, the avatar companion is always more impactful than the

agent companion. Additionally, the participants do not expect to be judged by an AI companion unlike by another person (Pickard et al., 2016). Considering that real-world AE was demonstrated even when the companion was not visually available to the participants and the monitoring was instead denoted by a light in the environment (Dumontheil et al., 2016), the companions' visual presence in the environment might not be as important for the AE as it is for the MPE. If AE requires a companion to have a subjective opinion, rather than just a visual presence, the agency of the companion (i.e., if it is driven by a human or an AI) should be crucial for the AE to elicit the SFE.

As argued in our previous study (Sutskova, Senju, & Smith., 2022), and as per TMSI (Blascovich et al., 2002), when measuring a virtual companion's social impact, both the social interaction context and what the participants know about the virtual companion should be considered. To the author's knowledge, there has not yet been a systematic study testing the SFE within IVR's, disentangling the contexts in which either the MPE or the AE elicits the SFE. The current experiment utilised the experimental control provided by IVR to dissociate the social companion's visual presence levels (not visible vs. humanoid vs. non-humanoid), within an immersive virtual space, from participants' interpretation of whether they were human-minded (avatar) vs. non-human minded (AI agent). We contrast companion attributes with the level of their attentiveness to the participant performing the cognitive task in an immersive space. By controlling whether the companion is monitoring or not monitoring the participant's performance, we were able to directly test whether the MPE or the AE explain the SFE within a virtual environment in a way not possible in a real-world scenario. The cognitive task utilised is the relational reasoning paradigm (RRP). The RRP reaction-times-based rapid response visual logic task engaging participant's executive functioning, prior shown to be susceptive to the AE in real-world testing (Dumontheil et al., 2016).

During the current experiment, the participants were immersed within a digitally generated 360-degree immersive virtual environment, whilst wearing a head-mounted display. Each participant performed the RRP within an immersive environment, whilst viewing the environment from a first-person perspective. During the RRP the participants are timed as they are matching rows of shapes and patterns performed at easy and difficult levels. The participants performed the RRP on a large virtual screen within the IVR, whilst in the presence of a companion or alone. The companion was attentive (or not attentive) to participant performance either with or without physical bodily presence (AE) and was merely bodily present (or absent) with or without attending to participants or their performance (MPE).

The companion's visual presence in the IVR was either visually absent (none), non-humanoid or humanoid. To control for possible confounds within the environment, such as distraction to the mere presence of any interactive non-social object, the participants in the visually absent companion group (none) performed the RRP task in the active presence of a non-social placeholder (an office fan) which mimicked the same movements, i.e., head turns towards the participant became rotations of the fan head. To test the impact of visually humanoid versus visually non-humanoid social presence, one group of participants performed in the presence of an autonomous video camera, the others in the presence of a humanoid virtual human. The companion's visually absent (none), non-humanoid and humanoid presence were entered into the analysis as *Companion Visual Presence* (CVP). All participants performed under the belief that they were within an IVR social context with either an AI (agent) or another person (avatar). Effects of the companion's capacity to mentalise (agency) were analysed under the factor of *Companion Agency*. The performance monitoring and non-monitoring (entered into the analysis as *Monitoring*) occurred orthogonally to the CVP and the Companion Agency. Within each level of the CVP, Monitoring occurred through either the camera (non-humanoid companion) or humanoid companion "eye". In the presence of the non-social object, the monitoring condition was indicated just through an on-screen instruction. To reduce the level of participants' anonymity in IVR, the participants were greeted in person by the researcher, who would operate the avatar to monitor the participant's performance, prior to the study. This manipulation was aimed at establishing that they are identifiable when they are being monitored performing. The participants

viewed the IVR environment from a first-person perspective. To reduce any self-avatar effects, such as identity occlusion and projection (Midha & Nandedkar, 2012; Peck et al., 2013), the participant was not given an avatar. Instead, all participants were notified that when performing, their behavioural markers (gaze, head rotation) and performance accuracy are seen by the observer, either AI or human (depending on the group assigned). This decision established that participants are aware that their actions are mapped onto their self-presence in IVE. Based on the conditions assigned, several predictions were made based on the MPE and the AE, the hypotheses and corresponding planned analyses are listed below.

Hypotheses and Planned Analyses

Canonical SFE is seen as an improvement in the easy task and a decrease in the difficult task under the social compared to the non-social condition. Note that the SFE may or may not show a canonical pattern, e.g., some prior studies only showed inhibition without facilitation (Emmerich & Masuch, 2016; Hoyt et al., 2003; Zanbaka et al., 2007). The figure below (**Figure 1**) illustrates a summary of the mixed design levels used in the current study. The predictions and corresponding MPE and AE contrasts within the omnibus mixed ANOVA design are listed below. For schematics of planned contrasts relating to the hypotheses see **Figure 2** for MPE and **Figure 3** for AE.

Design Summary

Between-Subjects	Between-Subjects	Within-Subjects	Within-Subjects
Companion Agency	**Companion Visual Presence**	**Monitoring**	**Difficulty**
AI Human	None Non-Humanoid Humanoid	Monitoring Not Monitoring	Easy Difficult

DV 1: Accuracy (%), **DV 2:** Reaction Times (ms) per accurate responses only
SFE: Better performance on easy and worse on difficult tasks, when in social context versus when not.

Figure 1: *A Summary Schematic of Independent Variables and Corresponding Levels in Experiment.*

Hypotheses based on the MPE. The first analysis tested the MPE within an immersive context, combining the companion visual presence (CVP) and task difficulty (CVP x Difficulty). There are two hypotheses based on the MPE, for an analysis plan related to the hypotheses of MPE see **Figure 2**.

The first hypothesis (*H1.a*) predicted that any companion presence irrespective of its visual type (CVP: non-humanoid or CVP: humanoid) would elicit the SFE, in contrast to when the participant is alone, i.e., companion visual presence is none (CVP: none), see **Figure 2, H1.a**

The second hypothesis (*H1.b*) predicted that the SFE would be affected more by the humanlike companion presence than the non-humanoid presence or no visual social presence (CVP: none), as per (Blascovich, 2002). The *H1.b* would be supported if the performance on easy conditions would improve, and performance on difficult conditions decreases, linearly as the CVP level of social influence increases from no companion presence (CVP: none), to CVP: non-humanoid, and with CVP: humanoid presence being of highest influence. The hypothesis was tested by contrasting the three levels of CVP (none, non-humanoid, humanoid) with one another. Similarly, to the first analysis, the effects are analysed at each level of task difficulty, see **Figure 2, H1.b.**
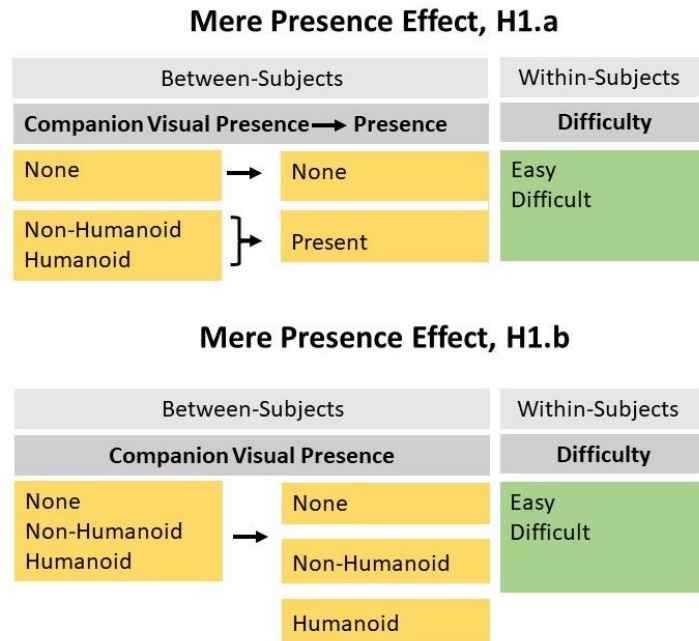
Mere Presence Effect, H1.a

Between-Subjects		Within-Subjects
Companion Visual Presence ➝ Presence		**Difficulty**
None	➝ None	Easy Difficult
Non-Humanoid Humanoid	}➝ Present	

Mere Presence Effect, H1.b

Between-Subjects		Within-Subjects
Companion Visual Presence		**Difficulty**
None Non-Humanoid Humanoid	➝ None	Easy Difficult
	Non-Humanoid	
	Humanoid	

Figure 2: *A Schematic of Analyses Plan for the Mere Presence Effect (MPE) hypotheses **H1.a** and **H1. b**.*

Hypotheses based on the AE. The second part of the analysis focused on the AE in an immersive environment, investigating whether the SFE in a co-immersive social environment can only be elicited through monitoring by a human-minded companion (Companion Agency: human) but not an AI-minded companion (Companion Agency: AI). To test this, the monitoring versus not monitoring condition was contrasted within each level of companion agency (AI, human) separately. There are two hypotheses based on the AE, for the analysis plan and related hypotheses of AE see **Figure 3**.
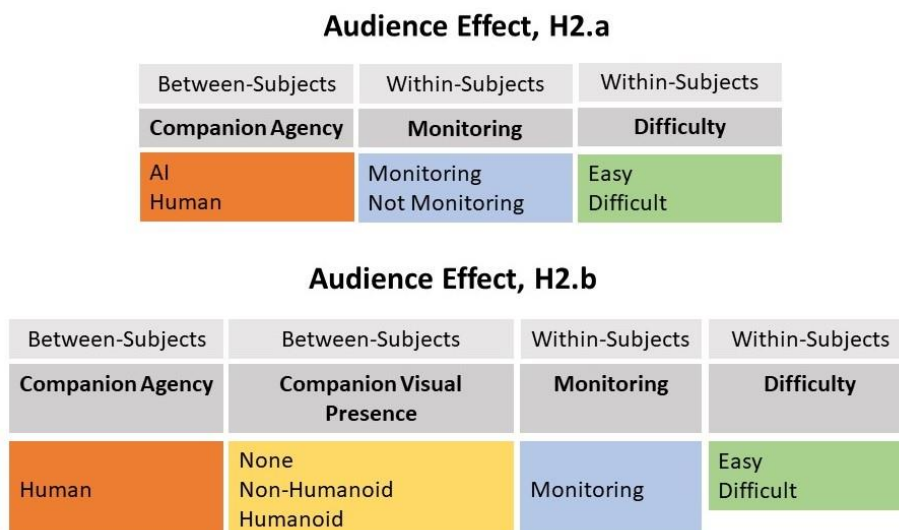
Audience Effect, H2.a

Between-Subjects	Within-Subjects	Within-Subjects
Companion Agency	**Monitoring**	**Difficulty**
AI Human	Monitoring Not Monitoring	Easy Difficult

Audience Effect, H2.b

Between-Subjects	Between-Subjects	Within-Subjects	Within-Subjects
Companion Agency	**Companion Visual Presence**	**Monitoring**	**Difficulty**
Human	None Non-Humanoid Humanoid	Monitoring	Easy Difficult

Figure 3: *A Schematic of Analyses Plan for the Audience Effect (AE) Hypotheses **H2.a** and **H2.b**.*

The first hypothesis, **H2.a,** predicts that participants' cognitive performance would change following the canonical pattern of the SFE when monitored versus when not, and the effect would be

present when the participants were allocated in Companion Agency: human condition, but not in Companion Agency: AI condition

Alongside the main predictions based on the AE focusing on Companion Agency, an additional analysis was conducted to test the effect of companion visual presence (CVP) on the AE in eliciting the SFE. The virtual interaction theories suggest that a higher human-likeness heuristic can "trick" the brain into processing a humanlike companion as more socially impactful (Nass et al., 1994; Nass & Moon, 2000) and that higher human-likeness might overall heighten the social impact even during human-minded agency of the companion (Blascovich et al., 2002). Therefore, the second hypothesis, **H2.b**, predicts that there might be a positive linear effect of the CVP, with higher levels of companion visual humanness, from no companion to non-humanoid to humanoid, which will increase the SFE. The effects were explored per difficulty and easy conditions separately.

Methods

Design

The experimental design included four factors, with two between-subjects factors, Companion Agency (AI, Human) and Companion Visual Presence (None, Non-Humanoid, Humanoid), and two within-subject factors, the level of Monitoring (Monitored, Not Monitored) under which participants performed, and task Difficulty (Easy, Difficult; see **Figure 1**). The Relational Reasoning Paradigm (RRP), which was also used in our recent study that measured SFE within a desktop-based online videoconferencing context (Sutskova et al., 2022), measured task performance in per cent accuracy and reaction times (RT) per accurate responses only. Participants who performed less than fifty per cent accuracy for easy and difficult conditions combined or performed more than three (3) standard deviations (SD's) away from the mean, were removed from the final analysis.

Participants

The data used for the final analysis consisted of data from 103 participants, 73 females and 30 males, with the age range of 18-55 ($M = 26.23$, $SD = 7.23$). A total of 138 participants were recruited to take part in the study, aiming to enter 18 participants per group for ANOVA analysis. The group sample size was powered (G*Power) at $1-\beta = .8$, $\alpha = .05$, Cohen $f = .44$, as per (Dumontheil et al., 2016) study, reporting real-world SFE with the current RRP paradigm, in Task 2 x Audience 2 interaction. A total of 35 participants were removed from the final data, for either not following the study instructions or due to their accuracy performance both on easy and difficult tasks combined falling under 50 per cent. Out of the 35, 19 participants were removed for not believing the experimental manipulation of monitoring or the companion agency (companion mind AI or human), see **Table 1** (**Results** section) for a breakdown of the remaining participants.

No prior experience in immersive virtual reality headsets was required to take part in the study. The simulation sickness risk was considered low, due to the stationary nature of the experiment and as a result of piloting the study on participants with a higher simulation sickness quotient (as measured by the Motion Sickness Susceptibility Questionnaire, MSSQ-Short (Golding, 2006), reporting no discomfort. The participants were reimbursed with either course credits or with £10 vouchers and fully debriefed on the social manipulation script after the completion of the experiment. The study was approved by Birkbeck, University of London, Ethics Committee: 181933.

Virtual Environment

The Immersive Virtual Environment (IVE) comprised of a brightly lit room with textured white walls, a desk with a chair, and a large flat TV screen on which the stimuli were presented (aerial virtual room perspective in **Figure 4, a**). The participants sat behind the virtual desk inside the environment, facing the TV screen positioned on the wall in front of the participant. In the real-world lab cubicle, the participant sat behind the computer wearing an Oculus HMD (**Figure 4, b**), pressing the keys on the keyboard, and responding yes (right arrow click) or no (left arrow click) on the RRP task.
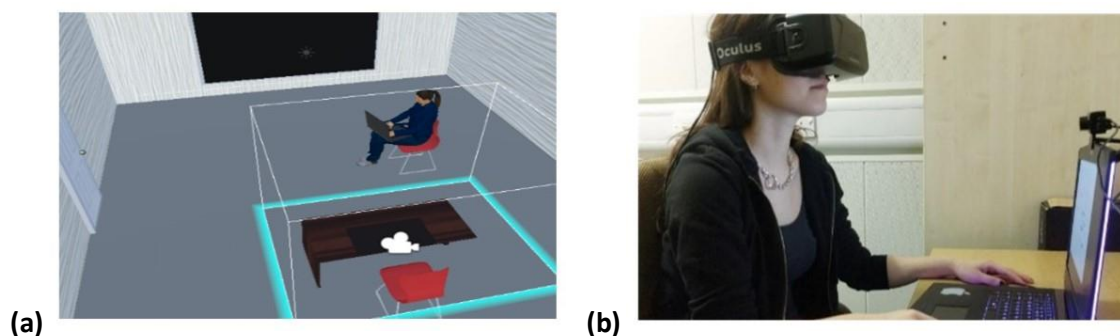
(a)   **(b)**

Figure 4: Virtual Environment Object Placement and Real-World Participant Positioning.

Note. **(a)** Top-down views of the virtual environment are set up from the perspective of the Unity game engine. The participants' desk was positioned at the back of the

virtual room. The camera icon highlights the participant's positioning behind the virtual table inside the IVE. **(b)** A participant taking part in the study wearing an Oculus Rift HMD using a virtual reality compatible laptop.

Depending on the experimental group assigned, the virtual companion was present either in a non-humanoid social presence form (observing interactive camera, **Figure 5, a**), in a humanoid form (an interactive humanoid character, **Figure 5, *b***), or as a visually non-present companion (companions animated presence replaced by an animated office fan, **Figure 5, c**). All objects were positioned ensuring the participants could fully observe the task on a virtual TV screen, but also notice any movement by the companion, or the non-social animated object (e.g., camera or fan), see **Figure 4, a**, for object placement within IVE (example with a humanoid companion present). During the training session, there was no companion implied, therefore the companion spot was empty, see **Figure 5, d.**
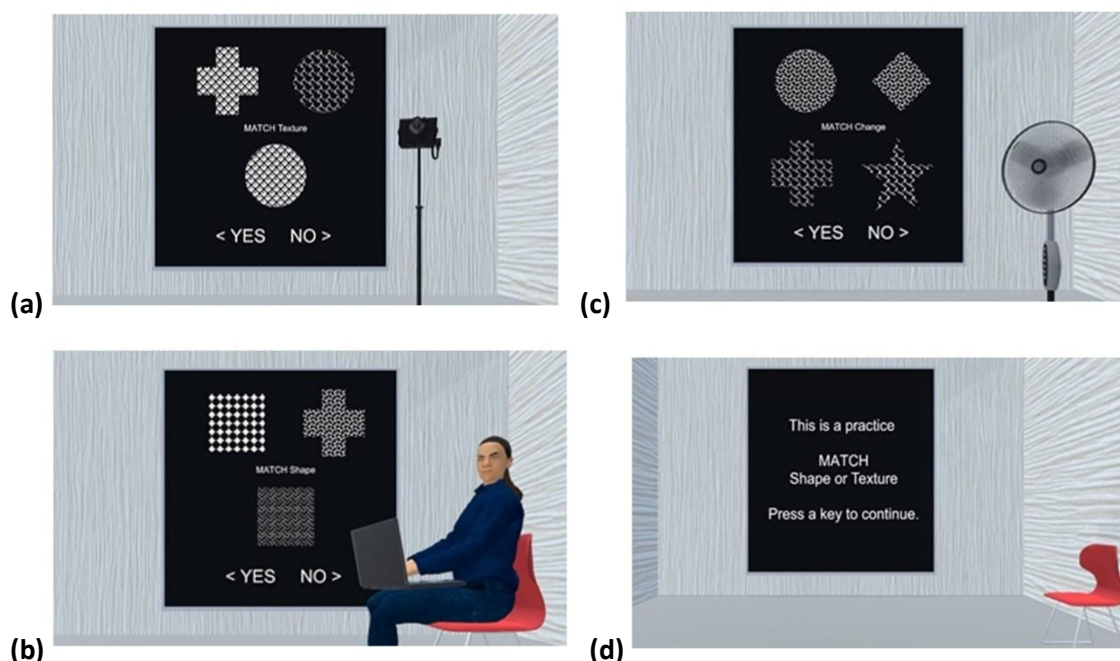

(a)   **(c)**

(b)   **(d)**

Figure 5: The Companion Monitoring Presence Conditions, from the Participant's Perspective Within the Immersive Space.

Note. The virtual environment setup from the participants' perspective, displaying the companions and the TV screen with the stimuli: **(a)** A present non-humanoid condition in which the participant

believed that they were monitored through an interactive controlled virtual camera (**b**); The humanoid presence condition, in which the participant believed they were monitored through the gaze of the humanoid character; **(c)** companion visually absent (none) condition, during which participants believed they could be monitored by either an AI or human, however without being able to see the companion. The **(d)** image shows an IVE layout of a practice session environment when no interactive objects were present or implied.

Social Context Instructions

Participants were led to believe that they were testing a new virtual monitoring and tracking software. Depending on the level of Companion Agency assigned to each participant, they were led to believe that they will be monitored in real-time by either an AI algorithm (AI condition) or a real human observer (Human condition), at some point during the task. The participants believed that their gazing behaviour (as projected through a VR headset) and performance were monitored by either another person (human mind) or an automated algorithm (AI processing). The companion monitoring blocks were marked by an IVE onscreen instruction "You are now being watched", which was then followed by the companion's monitoring behaviour as the participants performed the task. After the study's completion, all participants were asked whether they believed their companion was congruent with the manipulation. Contrary to the participant's belief, the social companion presence and monitoring behaviours within the IVE were all automated and matched across all the Companion Agency levels. There was no virtual real-time monitoring occurring at any time throughout the study by either companion. After the debriefing procedure revealed the social manipulation, the participants had a choice of withdrawing or committing their data to the analysis. Only the participants who consented to data inclusion and who believed in the social context script were included in the final analysis, nineteen participants were removed for not believing the social context manipulation (see *Participants*).

Companion's Social Presence and Monitoring

Depending on the level of CVP assigned to each participant, the companions' visual presence within IVE was either socially meaningful, i.e., a humanoid character (Humanoid condition), a non-humanoid camera (Non-Humanoid condition), or lacked socially meaningful presence, i.e., an animated office fan with no visible social presence value (None condition). All the participants experienced both the Monitored and Not Monitored conditions, which were allocated to different blocks and were initiated by instructions on the IVE TV task screen ("You are now being watched," You are now not being watched "). During the training session, the participants performed alongside a placeholder for the companion's visual presence in the study. The companion location during training sessions was substituted with a stationary chair (**Figure 5, d**) for all groups.

Out of three levels in the CVP, two were expected to be socially meaningful, the Non-Humanoid and Humanoid conditions. In the Non-Humanoid condition, the social companion presence was expressed in the form of an interactive camera on a tripod (**Figure 5, a**), which participants believed to be operated in real-time by either a monitoring real person (Companion Agency, Human) or by an autonomous AI algorithm (Companion Agency, AI). Participants believed that the companion, AI or human could monitor them inside the IVE using the camera (virtual CCTV). For the Humanoid condition, the companion was a humanoid virtual companion (**Figure 5, b**), believed to be controlled by the companion (Human or AI) in the real world. For both Non-Humanoid and Humanoid conditions, the participants believed that a real-world companion (AI, or human) could see into the IVE from the perspective of the virtual companion (either lens: non-humanoid camera, or gaze: humanoid).

During the Monitoring blocks, both the AI and human companion representations were animated to simulate monitoring motion, ranging between looking toward the participant and to the main task screen, as the participants performed. Both the non-humanoid and the humanoid companions gazing

motion was synchronised in a way that the head-turn towards and away from a participant in the humanoid character condition, would occur at the same time and for the same duration as the camera turn on the tripod, during all trials. In the Not Monitored blocks, the non-humanoid camera or the humanoid character's head tilted down and remained facing down throughout the condition block, with no observation motion directed towards the participant. To ensure a more natural not monitoring condition, in the case of humanoid character conditions, the character was looking down into their laptop.

By contrast, in the third level of the CVP (None condition), participants performed seemingly alone, with the only moving object in the immersive environment being an office fan. The participants believed that the office fan (**Figure 5, c**), experimentally used as a replacement for a social agent, was just part of the environment and operated independently from the monitoring conditions. There was no visually meaningful companion presence within the immersive space from the participants perspective. Without the participant's knowledge, in the companion visual presence: none condition, the office fan motion was matched to the companion visually present group's motion (camera, character) motion. The main difference between the non-social presence of a fan and the social presence of the non-humanoid and humanoid companion conditions was that the fan did not signal any companion social disposition towards the participant, therefore was socially meaningless.

Similarly, to the Non-Humanoid and Humanoid conditions, in the None condition, the participants were also made to believe that their performance is evaluated by either an AI algorithm or a human-minded companion, however, there was no visual representation of an interactive companion in the scene.

Task Stimuli

The current study used a timed visual logic task, the Relational Reasoning Paradigm (RRP), at two difficulty levels, easy and difficult. The task consisted of two rows of two grayscale images (**Figure 6**).
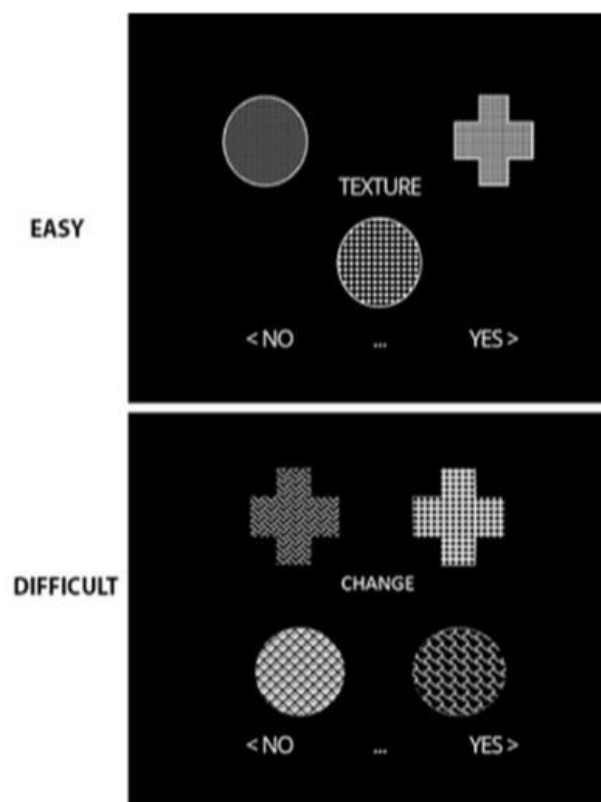

Figure 6: Illustration of The Relational Reasoning Paradigm Stimuli at Easy and Difficult Levels

For easy condition, the participants were asked to decide whether the top two images matched the bottom images in shape or pattern, by clicking the "yes" arrow if the bottom and top image matched and the "no" arrow if they did not. For the difficult task, the participants had to decide whether the images on the top changed in the same way (dimension: either shape or pattern) as the bottom images. By clicking the "yes" or "no" arrow, the participants matched the "change" type of the top and bottom rows. For example, if the top images change in shape, the bottom ones should also change in shape (not pattern), even if the shapes or patterns of the top and bottom are not the same. If both the top and bottom rows changed in shape then participants clicked "yes", if the top row changed in shape and the bottom row changed in the pattern the response would be "no".

Both easy and difficult conditions were presented for a fixed duration of 3.5 seconds, following 0.5 seconds of a blank screen, so participants had a total of 4 seconds to answer, until the transition to the next trial. If the answer was not given during the 4-second interval, the answer was considered incorrect. Each monitoring block (monitored, not monitored) consisted of two difficult and two easy sub-blocks, five trials per each sub-block, in total there were 20 trials per one monitoring condition, beginning with either a difficult or easy sub-block, followed by a complementing condition sub-block. Block sequences were counterbalanced between the groups and conditions. In total, there were eight monitoring blocks, four monitored and four not monitored counterbalanced. Every new monitoring block was preceded with a 10-second on-screen message on the IVE TV screen: "You are now being watched" or "You are now not being watched". The time of the message gave participants the time to look around and notice whether the companion was monitoring them or not (visually noticeable in the visually present companion groups). When the monitoring began, the virtual companion turned towards the participant and then at the screen, in no monitoring conditions the interactive virtual objects turned downwards, not facing the environment.

Procedure

Each participant was visually introduced to the RRP task outside of IVE and helped with adjusting the head-mounted display as they were introduced to the 360 environments within IVE. During the introduction phase, the participants were asked to perform a practice task. After the practice session was completed, and the participant reached a 70 per cent passing threshold both for Easy and Difficult trials, all participants were approached by the researcher, who made sure the participants were comfortable wearing the HMD and did not experience any simulation sickness.

The participants were assured that the researcher would be nearby in case of emergency, but unless the researcher would need to virtually monitor the participant's performance, they would be occupied otherwise. The participants were asked to inform the researcher when they completed the experiment, reinstating that otherwise, the researcher would not be able to see their performance. After the study completion, the researcher helped with removing the headset. The participants then filled in a form on whether they believed in monitoring by their assigned companion agency, and if so, whether they felt judged. The researcher made sure participants understood the questionnaire, guiding them through it when needed. Participants were fully debriefed after the study and made aware of the social scripting used for the experiment.

Apparatus

The virtual environment was developed in an open-source 3D platform Blender and imported into the Unity game engine. Stimuli presentation created and generated through the Unity platform. The RRP image textures (originally used in (Dumontheil et al., 2016) were edited to 20 % larger texture grain in Adobe Photoshop to reduce the Moire pattern effect inside the IVE. The humanoid character was a rigged 3D model mesh created in MakeHuman (www.makehumancommunity.org) free software and later imported to Blender for extra texturing. Both non-humanoid (tripod camera) and humanoid (digital researcher) were imported to and animated in the Unity game engine. The

experiment was conducted through a Unity game engine. Stimuli were presented on a virtual reality supporting Dell Alienware laptop and projected through the Oculus Rift DK2 (developers Kit) headset (resolution: 960 x 1080 pixels per eye, OLED display), refresh rate 75 Hz. The visible field of view is 93-degree horizontal, and 99-degree vertical, optical hardware is aspherical binocular lenses, IPD range of 63.3 mm fixed. The in-depth description of companion animation can be found in **Appendix A.**

Results

After excluding participants who did not meet inclusion criteria (see Participants section for details), there was an overall similar distribution of participant numbers in each of the six groups (allocated for each of three levels in the CVP factor for each of two levels in the Companion Agency factor).

Table 1: Number Of Participants Allocated to Each of Between Participant Conditions.

Companion Agency	Companion Visual Presence	Participants in Group
Human	Absent	18
	Present Non-Humanoid	18
	Present Humanoid	17
	Total	**53**
AI	Absent	17
	Present Non-Humanoid	17
	Present Humanoid	16
	Total	**50**
	Grand Total N	**103**

See **Table 1** for the number of participants allocated to each group. The data normality checks were conducted within each of the difficulty levels, for accuracy and RT separately. The test revealed that for accuracy, but not RT, the overall data distribution was significantly skewed as measured by Kolmogorov-Smirnov, for difficult t (103) = .161, p < .001 and easy t (103) = .144, p < .001. Therefore, the results for accuracy will be reported corrected when the data does not satisfy homogeneity or sphericity assumptions.

The summary tables of Means (M) and Standard Errors (**Table B2**) of accuracy percentage (**Table B1**) and reaction times per each factor and its corresponding levels are located in **Appendix B.**

Difficulty

For accuracy, there was a significant main effect of difficulty, F (1, 97) = 65.06, $p < .001$, ηp^2 = .40, power > 99%, with difficult trials (M = 86.87, SD = 11.30) performed significantly worse than easy (M = 94.24, SD = 4.78) trials. For RT, there was also a significant main effect of difficulty, F (1, 97) = 329.99, $p < .001$, ηp^2 = .77, power > 99%, with difficult trials (M = 2061.87, SD = 303.40) performed significantly slower than easy (M = 1640.22, SD = 266.24) trials.

H1: Mere Presence Effect (MPE)

As per the MPE prediction, companion visual presence (CVP) x difficulty interaction was significant in accuracy F (2, 97) = 4.69, p = .011, ηp^2 = .09 (Greenhouse-Geisser, power 77 %), and was marginal in RT, F (2, 97) = 2.39, p = .097, ηp^2 = .047 (Greenhouse-Geisser). The planned Bonferroni-corrected simple effects analyses investigated the interaction between the two levels of difficulty separately.

H1.a: Companion Presence vs Absence

The MPE hypothesis **H1.a** predicted that any companion visual presence (CVP: Humanoid and Non-Humanoid combined) would elicit the SFE when contrasted against the condition where a companion is absent (CVP: None).

Accuracy. As per the planned pairwise comparison, the interaction was broken down by difficulty, contrasting the impact of visually absent companion (CVP: None) versus visually present companion condition (CVP: Present; the Non-Humanoid and Humanoid conditions combined), at easy and difficult trials separately. The results showed no significant differences within the difficult or easy conditions. Difficult condition, t (100) = 0.72, p = .47 (absent M = 87.83, SD = 9.56, versus present, M = 86.37, SD = 12.13), easy condition, t (100) = 1.35, p = .18 (absent M = 95.09, SD = 3.7, versus present, M = 93.81, SD = 5.1).

Reaction Times. Similarly, to accuracy, the marginal CVP x difficulty interaction in the RT was broken down by difficulty, as planned. The results showed no significant difference in the planned contrast between the CVP presence and absence within difficult or easy conditions. Difficult tasks, t (100) = 0.81, p = .42 (absent M = 2096.45, SD = 230.78, versus present, M = 2044.04, SD = 334.89), easy tasks, t (100) = .90 p = .37 (absent M = 1609.40, SD = 238.30, versus present, M = 1656.10, SD = 279.90).

H1.b: Presence of Companion Type

The MPE hypothesis **H1.b** predicted that the visual presence of a humanoid companion (CVP: Humanoid) would be most impactful in contrast to the presence of CVP: Non-Humanoid and CVP: None. The impact was expected to increase linearly as social influence increases through companion visual characteristics, from None to Non-Humanoid to Humanoid).

Accuracy. The planned (Bonferroni corrected) pairwise comparisons tested the interaction at each level of CVP, broken down by levels of difficulty. For difficult conditions, the participants allocated to the Humanoid companion condition performed significantly worse than both those in Non-Humanoid and None conditions. Humanoid (M = 81.48, SD = 13.99) versus Non-Humanoid conditions (M = 90.99, SD = 7.79), p = .001, Humanoid versus None conditions (CVP: none) (M = 87.83, SD = 5.60), p = .048. Participants allocated to Non-Humanoid and None conditions did not statistically differ from one another, p = .67. For the easy conditions, the participants allocated to the Humanoid condition performed marginally worse than those in Non-Humanoid and None conditions. Humanoid condition (M = 92.53, SD = 5.67) versus Non-Humanoid (M = 95.02, SD = 4.45), p = .068, and None condition (M = 95.09, SD = 3.78), p = .058. The participants allocated to Non-Humanoid and None conditions did not statistically differ from each other (p > .99). There was no SFE facilitation of easy tasks, see **Figure 7, a.**
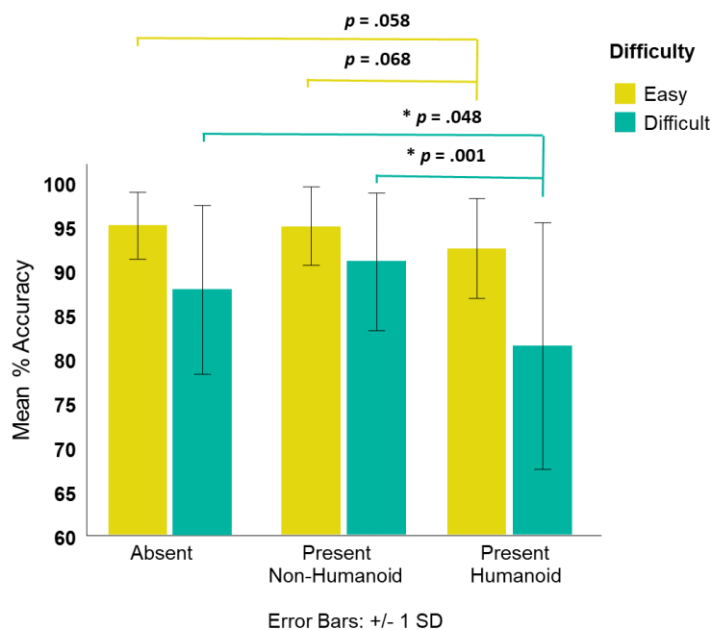
a. Accuracy
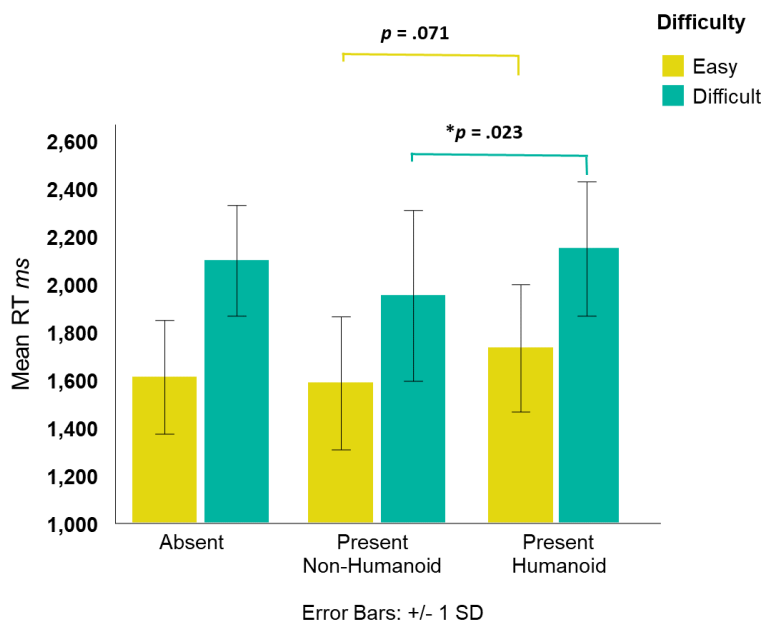
b. Reaction Times

Figure 7: MPE: Mean accuracies and reaction times (RT) at each level of the CVP and difficulty. Error bars: one Standard deviation (SD).

Reaction Times. The Bonferroni corrected pairwise comparisons explored the interaction at each level of CVP, broken down by levels of difficulty. For difficult conditions, the participants allocated to the Humanoid condition performed significantly slower than those in the Non-Humanoid condition but only numerically slower than those in the None condition. Humanoid (M = 2144.83, SD = 280.94)

versus Non-Humanoid (M = 1949.02, SD = 357.17), p = .023, and Humanoid versus None condition (M = 2096.54, SD = 230.79), p > .99. The participants allocated to Non-Humanoid and None conditions did not statistically differ from one another, p = .11. For easy conditions, the participants allocated to Humanoid condition performed marginally slower than those in Non-Humanoid condition but only numerically slower than those In None condition. Humanoid group (M = 1731.48, SD = 116.32) versus Non-Humanoid (M = 1584.98, SD = 277.33), p = .071, and Humanoid versus None conditions (M = 1609.39, SD = 238.30), p = .19. With no significant differences between participants allocated to None and Non-Humanoid conditions, p > .99. There was no significant facilitation see **Figure 7, b.**

MPE Summary. These results indicated that both RT and accuracy were negatively affected (inhibited) only by the presence of humanoid companion (CVP: Humanoid), both on easy and difficult tasks (as per ***H1.b***.), but not by the presence of a Non-Humanoid companion, not supporting MPE ***H1.a***. There was no significant linear effect as companion's visual humanness increases, with no significant difference between the presence of non-humanoid (CVP: Non-Humanoid) and absent companion (CVP: None). Although, the impact of humanoid presence did elicit a social response, in contrast to other types of presence, as per ***H2.b***. The effects were inhibitory, without facilitation, which is not in line with the canonical pattern of the SFE as hypothesised, but in line with the majority of IVR findings reported to date.

H2: Audience Effect (AE)

The AE (***H2.a***) hypothesis predicted that there would be a significant Companion Agency x Monitoring x Difficulty interaction. The Monitoring should influence participants' performance in human-minded Companion Agency: Human group, but not Companion Agency: AI. The effects should be irrespective of CVP. The interaction between Companion Agency, Monitoring and Difficulty was non-significant *F (1, 97) = 0.28, p = .87, ηp^2 < .001* for accuracy, and in RT *F* (1, 97) = .526, *p* = .470, *ηp^2* = 0.005, not supporting our hypothesis. However, as planned, the *interaction was broken down, contrasting Monitoring levels within each level of Companion Agency and Difficulty (see **H2.a)**.*

In ***H2.b*** we predicted that the companion's visual presence and humanoid form could contribute additionally to the impact of the monitoring companion, predicting a Companion Agency x Monitoring x CVP x Difficulty interaction. The interaction was not significant for accuracy *F* (2, 97) = 2.14, *p* = .12, *ηp^2* = 0.042, or for RT *F* (2, 97) = 1.20, *p* = .25, *ηp^2* = 0.028. However, as planned, the interaction was broken down, contrasting performance outcomes under different CVP levels during Monitoring. The effects were contrasted within the human-minded Companion Agency: Human group, per each Difficulty level separately. There was no significant effect for the AI-companion group, therefore CVP x monitoring interaction was not followed up (see ***H2.b***).

H2.a: AE Irrespective of CVP

Accuracy. The follow-up analysis for the main effect of monitoring explored whether the performance changes within each level of the Companion Agency (Human or AI) were different from each other. The results showed that although the monitoring decreased participants' performance accuracy numerically, both in the AI group (not monitored M = 89.49, SD = 9.09, versus monitored M = 89.15, SD = 8.11), and in the Human group (not monitored M = 92.47, SD = 6.21, versus monitored M = 90.98, SD = 6.39), the effect was only significant in the participants allocated to Human Companion Agency condition (p = .021, ηp^2 = .054) but not those allocated to AI condition (p = .62, ηp^2 = .003). The planned follow-up analysis of Monitoring x Difficulty interaction at each level of Companion Agency revealed that monitoring by a human companion marginally decreased the accuracy of the performance on the difficult trials (p = .067, ηp^2 = .034), with no significant changes on easy trials (p = .139, ηp^2 = .005), see **Figure 8, a**. By contrast, monitoring by AI companion did not significantly change performance either in easy trials p = .478, ηp^2 = .022 or in difficult trials p = .894, ηp^2 < .001.
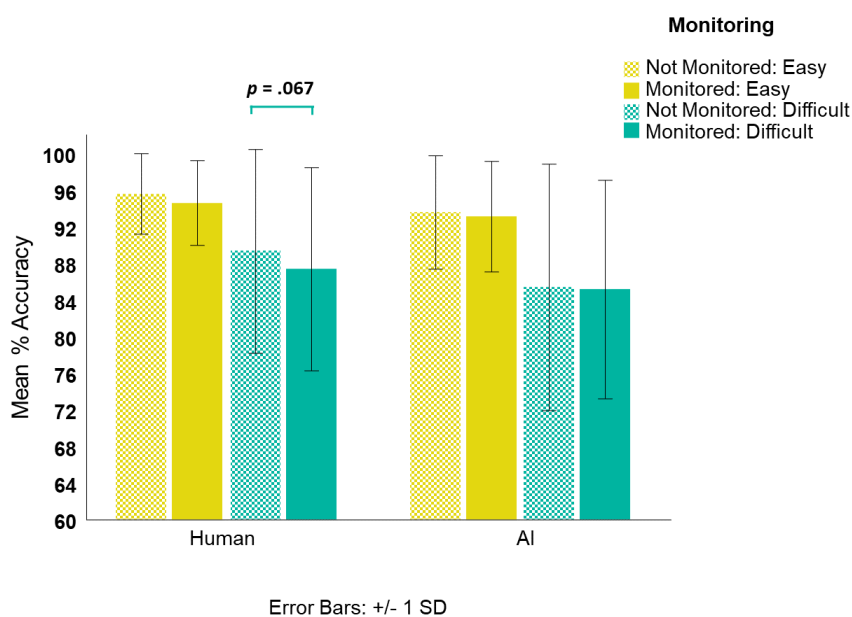
a. Accuracy
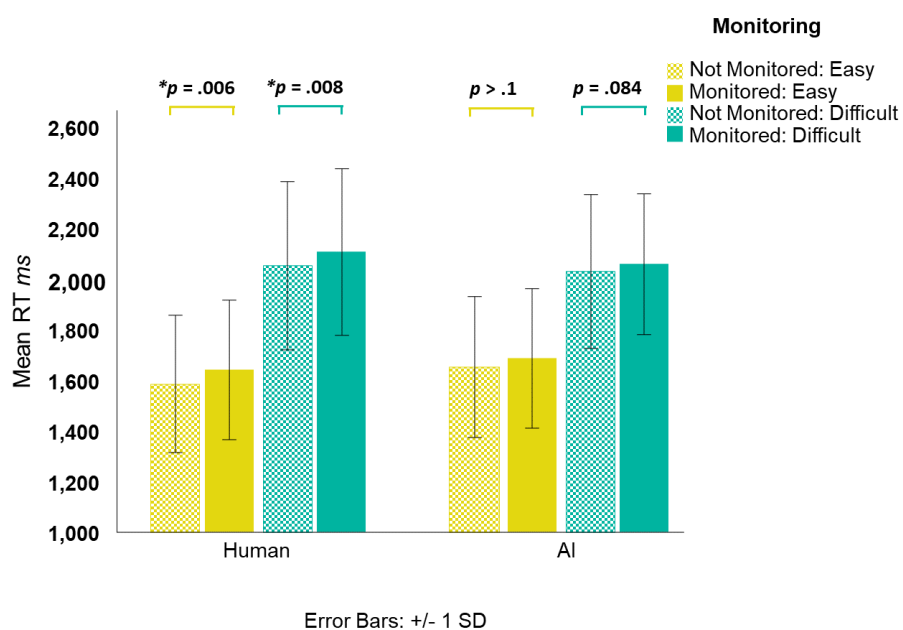
b. Reaction Times

Figure 8: AE: Irrespective of CVP, contrasting Monitored versus Non-Monitored conditions within Human and AI companion groups separately, at Easy and Difficult Levels of Performance. Descriptive Means and One Standard Deviation (SD) Per Each Condition.

Reaction Times. For RT, the results showed that participants allocated to both Companion Agency conditions performed slower when monitored (Human, $M = 1873,08$, $SD = 274.27$; AI, $M = 1872.57$, $SD = 247.53$), versus not monitored (Human $M = 1818.64$, $SD = 279.70$, AI $M = 1840.86$, $SD = 256.66$).

However, the effect was larger in those allocated to the Human condition $p < .001$, $\eta p^2 = .125$, than in those allocated to the AI condition, $p = .040$, $\eta p^2 = .043$. The planned follow-up analysis breaking down the effect further by each level of Difficulty (see **Figure 8**, **b**), there was a marginal performance decrease when monitored for easy trials (marginal, $p = .084$, $\eta p^2 = .030$) but none for difficult trials ($p = .181$, $\eta p^2 = .018$) for participants allocated to AI condition. For those allocated to the Human condition, by contrast, the monitoring significantly affected RTs in both the difficulty levels, with performance significantly slower both for easy, $p = .006$, $\eta p^2 = .075$, and difficult trials, $p = .008$, $\eta p^2 = .071$, when monitored.

H2.b: AE accounting for CVP

The planned Bonferroni corrected contrasts between the levels of Monitoring for each level of Companion Agency, CVP and Difficulty, revealed that for participants allocated to the Human Companion Agency condition, the monitoring by the humanoid CVP decreased performance on difficult tasks in contrast to non-humanoid companion. The effect was significant for accuracy, Humanoid ($M = 81.90$, $SD = 14.43$) versus Non-Humanoid ($M = 91.20$, $SD = 5.71$), $p = .048$, $\eta p^2 = 0.064$, and marginal for RT, Humanoid ($M = 2237.10$, $SD = 356.73$) versus Non-Humanoid ($M = 2011.50$, $SD = 346.68$), $p = .081$, $\eta p^2 = 0.050$. There we no other significant differences between CVP levels when monitored.

AE Summary

Based on the results from breaking down each companion agency level of the Monitoring x Difficulty interaction, monitoring decreased performance overall both for human (significant) and AI (marginally) companions. However, when broken down by difficulty (see **Figure 8**), the results support *H2.a* TMSI theory, suggesting that only a human-minded companion would affect performance during monitoring. The canonical SFE-related facilitation was not observed, instead, we observed overall performance inhibition when monitored.

The ***H2.b,*** stating that a companion's humanlike visual form facilitates AE was not supported, however, the human visual form of the monitoring companion significantly decreased the participant's performance in contrast to the non-humanoid monitoring companion's presence.

Companion Visual Presence (CVP)

There was an overall significant main effect of CVP in accuracy, $F(2, 97) = 7.22$, $p = .001$, $\eta p^2 = .13$, and RT, $F(1, 97) = 3.88$, $p = .024$, $\eta p^2 = .074$. The effects were explored in post-hoc independently from the SFE effect, such as performance difficulty. The Bonferroni corrected pairwise contrasts tested the performance differences under the three different CVP levels.

CVP Type Impact

Accuracy. Contrast revealed that a Humanoid companion ($M = 87.00$, $SD = 8.55$) led to significantly worse performance than both the Non-Humanoid companion ($M = 93.0$, $SD = 5.61$), $p = .001$, and None condition ($M = 91.46$, $SD = 6.02$), $p = .024$. Accuracies did not differ between Non-Humanoid and Humanoid conditions, $p > .99$.

Reaction Times. The performance was slowest in the Humanoid condition ($M = 1922.19$, $SD = 242.45$), being significantly different to the Non-Humanoid condition ($M = 1762.20$, $SD = 288.27$), $p = .019$. The difference between Humanoid and None conditions was only numerical and not significant ($M = 1842.59$, $SD = 204.10$, $p = .55$). The Non-Humanoid condition showed numerically fastest RTS overall, however not statistically different from the next fastest, None condition ($p = .44$).

Overall, both for accuracy and RT, the results are in line with findings in the *MPE: Presence of Companion Type* above, suggesting that when performing under a humanoid companion, performance was significantly less accurate and slower, irrespective of task difficulty.

Monitoring

Looking into data-driven effects, there was also a significant main effect of monitoring, both in accuracy $F(1, 97) = 3.95$, $p = .050$, $\eta p^2 = .039$ and RT $F(1, 97) = 16.66$, $p < .001$, $\eta p^2 = .147$. The post-hoc analyses revealed that for accuracy, the Not Monitored condition ($M = 91.03$, $SD = 7.85$) showed overall more accurate performance than the Monitored condition ($M = 90.09$, $SD = 7.29$). For RT, the Monitored ($M = 1872.83$, $SD = 267.67$) conditions performed significantly slower (worse) than the Not Monitored condition ($M = 1829.43$, $SD = 267.67$). Overall, the process of monitoring, irrespective of companion agency, was also significantly detrimental to performance. However, when breaking down effects by Monitoring the effect is only significant in human-minded companion conditions.

Discussion

The current study investigated the social facilitation effect (SFE), within an immersive virtual environment, focusing on comparing the two hypothesised mechanisms of eliciting SFE, the audience effect (AE) and the mere presence effect (MPE). The experiment measured participant's cognitive performance (on a relational reasoning paradigm: RRP) changes under the mere co-presence with the companion at three levels of companion visual presence (CVP: none, non-humanoid, humanoid) – MPE, and when the companion was monitoring the participant's performance – AE.

Firstly, the MPE hypothesis *H1.a* predicted that any social virtual companion presence (CVP: non-humanoid, and CVP: humanoid combined) in an immersive environment will elicit SFE, in contrast to when a companion is not visually co-present with the participant (CVP: none). This hypothesis was not supported by the data. However, hypothesis *H1.b*, that the companion humanoid immersive presence would be most impactful, was indeed supported by the results. Tasks performed under humanoid companion demonstrated a significantly worse outcome than in other CVP conditions, with no significant differences between non-humanoid and no companion (CVP: none) groups. Although the effect of humanoid versus other CVP conditions was significant for accuracy and reaction times (RT), the results showed an overall performance inhibition and not facilitation, not replicating canonical SFE (facilitation in easy tasks and inhibition in difficult tasks) observed in real-world interaction but in line with the majority of recent immersive SFE literature (Emmerich & Masuch, 2016; Hoyt et al., 2003; Zanbaka et al., 2007). Based on the current virtual MPE findings, the humanoid presence was possibly socially distracting to the participants, with no evidence of positive social facilitation in easy tasks. This conclusion could be supported by a significant overall main effect of CVP for accuracy and RT, decreasing participants' performance when a humanoid companion was present irrespective of difficulty. Future studies recording eye tracking during this task could shed more light on whether humanoid companions are more distracting than non-humanoid companions. For example, eye tracking could be used to investigate whether the participants' gaze disengages from the task, towards the companion, measuring the frequency of gaze saccades towards the companion, leading to worse task performance. However, as mentioned by (Guerin, 1983) people often restrict their overt embodied behaviours, such as body movement, during co-presence, and potentially inhibit gaze towards the companion due to civil attentiveness. Therefore, a more multimodal approach could be taken, transcribing participants' behaviours in the context of different companions and virtual interaction conditions. Similarly, to Klowait (2023) and (Klowait & Erofeeva, 2023), the analysis could focus on more general non-verbal behaviours, such as sequential analysis of body motion and action-based decision-making. A systematic analysis of such behaviours could reveal more about inner processes related to perceived others' presence, which might not be revealed through focusing just on eye-tracking.

For the AE, there were two hypotheses, with the planned comparison of the companion agency × monitoring × difficulty interaction. For AE, *H2.a*, predicted that monitoring would only impact human-operated companion groups, but not the groups with AI-operated companion groups. The results showed that indeed the performance was significantly affected overall only during monitoring by another person (human-mind). For accuracy, the effect was overall detrimental to performance,

with no facilitation, marginally in difficult tasks, and significantly detrimental both in easy and difficult tasks in RT. There was also an overall significant effect of monitoring, demonstrating that being monitored in IVE, irrespective of companion agency, can be significantly detrimental to participants' cognitive performance, both in RT and accuracy. However, as found in *H2.a.* the results were driven by significant effects in human-minded companions but not AI groups. The results overall support the AE hypothesis *H2.a*, that monitoring by a human companion, but not AI, impacts participants' performance significantly. As with MPE, there was no SFE-related facilitation.

Additionally, the analysis investigated the contribution of increasing social impact to AE, through increasing companion visual presence (CVP), from none to non-humanoid to humanoid. *H2.b* predicted that the companion visual presence might amplify the effect of monitoring, suggesting a linear positive relationship of performance impact as the CVP increases from none to non-humanoid to humanoid. The test was conducted only in the human-minded companion group in which monitoring resulted in the predicted significant effect irrespective of CVP (*H2.a).* When testing differences between each level of CVP while being monitored, the humanoid companion group performed significantly less accurately and marginally slower than the non-humanoid companion group. However, there was no linear relationship between the increasing impact of CVP. Therefore, the results did not support the hypothesis (*H2.b*) that increasing higher levels of virtual companions' social presence contributes additionally to AE impact. However, importantly, the presence of monitoring humanoid companion did decrease performance in addition to non-humanoid companion presence. This could suggest that even for AE, the humanoid features of companions in IVR are important.

In summary, we found that sharing the same immersive environment with a humanoid companion (MPE), irrespective of the companion's agency (human, AI), influenced participants' cognitive performance outcomes (SFE). We also found that the belief of being monitored (AE) by another person, not AI, influenced participants' performance significantly, irrespective of whether the virtual observer was visible. Therefore, the socially motivated cognitive performance change (SFE) during immersive virtual interaction can be influenced both through visual aspects of the virtual companion in the shared space, and when the social evaluation might take place – participants' beliefs about whether the companion had a human mind. Additionally, as observed in the results of analyses based on *H2.b,* there might be an accumulative effect of humanlike visual attributes companion and the belief that monitoring occurs by a human-minded companion who is able to mentalise. Similar accumulative effects were found in our videoconference-based equivalent of the current study (Sutskova et al., 2022). Future research needs to test this notion further, both in IVR and other virtual settings.

The current immersive experiment findings were overall inhibitory, both when testing humanoid companions' mere co-presence and human-minded companion's monitoring. We did not find a canonical interaction, facilitation in easy tasks and inhibition in difficult tasks, often attributed to SFE (Bond & Titus, 1983; Guerin & Innes, 1984; Uziel, 2007). Interestingly, the pattern of social inhibition without facilitation was the most reported effect in the current immersive SFE studies. Conclusively, it seems that immersive companions are significantly socially influential, yet the impact does not seem to improve the participant's performance on cognitive tasks just detriments it.

One possible explanation for the lack of facilitation could be that, as mentioned in the SFE meta-analysis, even in the in-real world face-to-face SFE research the results on the easy conditions are often harder to replicate (Bond & Titus, 1983). Considering the easy task was often at the ceiling during the study, it is possible that participants had little room for improvement for significant facilitation to be noticed. To further investigate these findings, future research needs to apply additional physiological methods to accompany the current behavioural IVR paradigm, and possibly use a more taxing cognitive task.

There are several possible explanations for the lack of facilitation in our findings. Firstly, the immersive environment is considered to be overall more cognitively taxing, even in contrast to the video conference-based interaction in the real world (Makransky et al., 2019). This could mean that the higher overall cognitive load imposed by the IVR could have increased the overall difficulty level of performing the cognitive tasks. As per SFE, the difficult task performance gets inhibited rather than facilitated during the social context in contrast to performing alone. If both the easy and difficult tasks were considered challenging under a higher cognitive load within the IVR environment, then adding additional social influence could have led to an inhibitory effect due to cognitive overwhelming. However, when looking at the overall mean accuracy of performance, the performance on easy tasks was almost at the ceiling, suggesting that cognitive load does not fully explain the inhibitory trend of being overwhelmed.

Alternatively, it is possible that there was no facilitation to be found in an immersive space because there was no sufficient main driving factor for facilitation. As discussed in the introduction, realistic self-representation, and a sense of self-presence within the immersive environment, might be one of the crucial factors which lead to pro-social motivation when monitored (AE) or vigilance when merely co-present (MPE). The realistic self-presence of participants in IVR in the current study was limited mainly due to the hardware constraints of the virtual head-mounted display at the time of testing. Although we attempted to elicit a higher sense of self-presence through participants believing that their gaze and performance can be seen and they are identifiable as they perform, this level of functional presence is likely not sufficient for accountability and therefore pro-social action. Indeed, in our virtual desktop study, where the participant was remotely video present with companions, the SFE were mostly facilitatory (Sutskova et al., 2022). Without the self-present facilitation, the inhibitory presence of co-immersive humanoid companions could have been more socially distracting than the other CVP presence types. The effect however is not necessarily unique to second-person cognition (Schilbach et al., 2013), as required for SFE.

Additionally, considering that the sense of self-presence in the environment is important, it is possible that when the participants were wearing the headset, the presence of the researcher in the real-world testing cubicle could have already raised the levels of participants' social arousal. Even though the participants were shown that during testing the researcher is sitting with their back to the participant, working on their own project, the participants could have been more vigilant in the real-world presence of another person, as their own view of the real-world environment was restricted by the virtual headset. The initial real-world co-presence could have contributed to the virtual social influence of monitoring or humanoid presence, overall raising the arousal levels high enough to detriment the performance. This could explain the overall high level of initial performance, even in the alone condition, then dropping when additional virtual social influence was introduced. If this is indeed the case, then immersive co-presence and monitoring can be impactful as per SFE, however additional real-world environment can accumulatively contribute to the overall arousal, rendering it detrimental. The additive social effects of real-world and immersive environments are an interesting concept and should be tested. If this is indeed the case, future interventions using IVR should be mindful also of the physical surroundings.

In the present study, the only sensory modality used to induce co-presence was vision. It is important to however note the important role the multimodal cues can play to invoking a sense of immersion and social presence, including proprioception, haptics (touch), olfactory (smell) and auditory cues (Klowait, 2023; Klowait and Erofeeva, 2021), but were beyond the scope of present study. Any sensory input in addition to vision, or an expectation of such, could additionally contribute to either the congruency or incongruency of the immersive experience, as well as the vigilance over the occluded physical environment in which immersion occurs. Future studies should therefore focus on the multisensory experience of co-immersion. Indeed, in their review article, Martin and colleagues (2022) found that several industries using immersive technologies are already relying on other sensory modalities in their IVE to elicit higher levels of realism and immersion. It is possible that additional

social sensory inputs, such as voice or sound of breathing, as well as affordance of touch in IVE, could elicit higher levels of immersive co-presence. Reducing the same sensory affordances from the physical surroundings could contribute to a lesser focus on the world outside of IVE. Additionally, the individual differences in social processing and personality traits can also influence SFE (Uziel, 2007), therefore the future work should consider such differences when designing experimental controls and utilising SFE in IVE practical implementations.

As social media companies and technology developers are currently shaping the future of augmented social interaction, there are already new opportunities that enable testing levels of self- and companion presence, mixing both virtual and own environment. The virtual CAVEs and augmented reality can bring virtual into the real-world environment, without compromising on both. The realistic real-time face scanning and digital twins can recreate and translate participants' realistic replicas to immersive and augmented spaces. All of these new developments create a mixed reality in which the self-presence of the protagonist can be realistic at different levels alongside projected or immersive companions. The SFE research into augmented reality has already shown promising results through virtually projecting the companions into participants physical environments (Mostajeran et al., 2022).

All of these new platforms and methods will undoubtedly enable cognitive and social perception testing at levels previously impossible in real-world communication, broadening the understanding of the human brain and what it means to be social. It is however important to understand the constraints of each emerging technology.

Although the current experiments' findings are interesting, more questions arise through these immersive results and their generalisation to the in-real-world social scenarios. The immersive environments are no doubt robust platforms for testing the cognitive impacts of social interaction with others. Paraphrasing Blascovich and colleagues (2002), immersive virtual reality is a unique tool that helps us to pick apart and reverse engineer the most complex behaviours in a controlled systematic, yet mostly ecologically valid way.

It is important to note however that, as with every technological tool, there are limitations that need to be considered, and immersive virtual interaction is currently far from replicating real-world communication with high validity. Establishing congruency and balance between all the sensory modalities in a realistic way is still a work in progress. Irrespective, immersive reality seems to be the right tool to explore some of the emerging trends of virtual and mixed social interaction. It is important to continue immersive research, especially considering the interest in social immersive experiences increases, and these technologies improve rapidly reaching levels of what was considered science-fiction just a few decades ago.

Data Availability

The means and standard errors or accuracy (%) and reaction times (ms) of all the levels of the independent factors presented in the current study are available in **Appendix B.**

Conflict of Interest

The authors of the current manuscript declare that the current research was conducted in the absence of financial influences that might have led to the perceived or potential conflict of interest.

References

Appel, J., Von Der Pütten, A., Krämer, N. C., & Gratch, J. (2012). Does humanity matter? analyzing the importance of social cues and perceived agency of a computer system for the emergence of social reactions during human-computer interaction. *Advances in Human-Computer Interaction*, *2012*. https://doi.org/10.1155/2012/324694

Blascovich, J. (2002). *A Theoretical Model of Social Influence for Increasing the Utility of Collaborative Virtual Environments*.

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive Virtual Environment Technology as a Methodological Tool for Social Psychology. *Psychological Inquiry*, *13*(2), 103–124. https://doi.org/10.1207/S15327965PLI1302_01

Bond, C. F., & Titus, L. J. (1983). Social facilitation: A meta-analysis of 241 studies. *Psychological Bulletin*, *94*(2), 265–292. https://doi.org/10.1037/0033-2909.94.2.265

Boothby, E. J., Smith, L. K., Clark, M. S., & Bargh, J. A. (2017). The world looks better together: How close others enhance our visual experiences. *Personal Relationships*, *24*(3), 694–714. https://doi.org/10.1111/PERE.12201

Dumontheil, I., Wolf, L. K., & Blakemore, S.-J. (2016). Audience effects on the neural correlates of relational reasoning in adolescence. *Neuropsychologia*, *87*, 85–95. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2016.05.001

Emmerich, K., & Masuch, M. (2016). The influence of virtual agents on player experience and performance. *CHI PLAY 2016 - Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, 10–21. https://doi.org/10.1145/2967934.2968092

Eskinazi, M., & Giannopulu, I. (2021). Continuity in intuition and insight: from real to naturalistic virtual environment. *Scientific Reports 2021 11:1*, *11*(1), 1–12. https://doi.org/10.1038/s41598-021-81532-w

Gasteiger, N., Loveys, K., Law, M., & Broadbent, E. (2021). Friends from the Future: A Scoping Review of Research into Robots and Computer Agents to Combat Loneliness in Older People. *Clinical Interventions in Aging*, *16*, 941. https://doi.org/10.2147/CIA.S282709

Golding, J. F. (2006). Predicting individual differences in motion sickness susceptibility by questionnaire. *Personality and Individual Differences*, *41*(2), 237–248. https://doi.org/10.1016/j.paid.2006.01.012

Green, R. G., & Bushman, B. J. (1989). Presence. In H. Wagner & A. Manstead (Eds.), *Handbook of Psychophysiology* (H. Wagner, pp. 261–281).

Guerin, B. (1983). Social facilitation and social monitoring: A test of three models. *British Journal of Social Psychology*, *22*(3), 203–214. https://doi.org/10.1111/J.2044-8309.1983.TB00585.X

Guerin, B. (1986). Mere presence effects in humans: A review. *Journal of Experimental Social Psychology*, *22*(1), 38–77. https://doi.org/10.1016/0022-1031(86)90040-5

Guerin, B., & Innes, J. M. (1984). Explanations of social facilitation: A review. In *Current Psychology* (Vol. 3, Issue 2, pp. 32–52). Springer-Verlag. https://doi.org/10.1007/BF02686548

Hamilton, A. F. de C., & Lind, F. (2016). Audience effects: what can they tell us about social neuroscience, theory of mind and autism? *Culture and Brain*, *4*(2), 159–177. https://doi.org/10.1007/s40167-016-0044-5

Hayes, A. L., Ulinski, A. C., & Hodges, L. F. (2010). That avatar is looking at me! Social inhibition in virtual worlds. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6356 LNAI*, 454–467. https://doi.org/10.1007/978-3-642-15892-6_49

Hoyt, C. L., Blascovich, J., & Swinth, K. R. (2003). Social Inhibition in Immersive Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *12*(2), 183–195. https://doi.org/10.1162/105474603321640932

Huang, M. H., & Rust, R. T. (2018). Artificial Intelligence in Service. *Journal of Service Research*, *21*(2), 155–172. https://doi.org/10.1177/1094670517752459

Iqbal, M. J., Javed, Z., Sadia, H., Qureshi, I. A., Irshad, A., Ahmed, R., Malik, K., Raza, S., Abbas, A., Pezzani, R., & Sharifi-Rad, J. (2021). Clinical applications of artificial intelligence and machine learning in cancer diagnosis: looking into the future. *Cancer Cell International*, *21*(1), 270. https://doi.org/10.1186/s12935-021-01981-1

Klowait, N. (2023). On the Multimodal Resolution of a Search Sequence in Virtual Reality. *Human Behavior and Emerging Technologies*, *2023*. https://doi.org/10.1155/2023/8417012

Klowait, N., & Erofeeva, M. A. (2021). The Rise of Interactional Multimodality in Human-Computer Interaction. *Monitoring Obshchestvennogo Mneniya: Ekonomicheskie i Sotsial'nye Peremeny*, *1*, 46–70. https://doi.org/10.14515/MONITORING.2021.1.1793

Klowait, N. O., & Erofeeva, M. (2023). Halting the Decay of Talk: How Atypical Interactants Adapt their Virtual Worlds. *Social Interaction. Video-Based Studies of Human Sociality*, *6*(1). https://doi.org/10.7146/si.v6i1.136903

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, *60*, 225–236. https://doi.org/10.1016/J.LEARNINSTRUC.2017.12.007

Martin, D., Malpica, S., Gutierrez, D., Masia, B., & Serrano, A. (2022). Multimodality in VR: A Survey. *ACM Computing Surveys (CSUR)*, *54*(10 s). https://doi.org/10.1145/3508361

Midha, V., & Nandedkar, A. (2012). Impact of similarity between avatar and their users on their perceived identifiability: Evidence from virtual teams in Second Life platform. *Computers in Human Behavior*, *28*(3), 929–932. https://doi.org/10.1016/j.chb.2011.12.013

Mostajeran, F., Reisewitz, P., & Steinicke, F. (2022). *Social Facilitation and Inhibition in Augmented Reality: Performing Motor and Cognitive Tasks in the Presence of a Virtual Agent; Social Facilitation and Inhibition in Augmented Reality: Performing Motor and Cognitive Tasks in the Presence of a Virtual Ag*. https://doi.org/10.1109/VRW55335.2022.00073

Nass, C., & Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues*, *56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Nass, C., Steuer, J., Henriksen, L., & Dryer, D. C. (1994). Machines, social attributions, and ethopoeia: performance assessments of computers subsequent to "self-" or "other-" evaluations. *International Journal of Human-Computer Studies*, *40*(3), 543–559. https://doi.org/10.1006/IJHC.1994.1025

Oh, C. S., Bailenson, J. N., & Welch, G. F. (2018). A Systematic Review of Social Presence: Definition, Antecedents, and Implications. *Frontiers in Robotics and AI*, *5*, 114. https://doi.org/10.3389/frobt.2018.00114

Olk, B., Dinu, A., Zielinski, D. J., & Kopper, R. (2018). Measuring visual search and distraction in immersive virtual reality. *Royal Society Open Science*, *5*(5). https://doi.org/10.1098/RSOS.172331

Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, *22*(3), 779–787. https://doi.org/10.1016/j.concog.2013.04.016

Pickard, M. D., Roster, C. A., & Chen, Y. (2016). Revealing sensitive information in personal interviews: Is self-disclosure easier with humans or avatars and under what conditions? *Computers in Human Behavior*, *65*, 23–30. https://doi.org/10.1016/j.chb.2016.08.004

Rajecki, D. W. (2010). Zajonc, Cockroaches, and Chickens, c. 1965—1975: A Characterization and Contextualization. *Emotion Review*, *2*(4), 320–328. https://doi.org/10.1177/1754073910375477

Rajecki, D. W., Ickes, W., Corcoran, C., & Lenerz, K. (1977). Social Facilitation of Human Performance: Mere Presence Effects. *The Journal of Social Psychology*, *102*(2), 297–310. https://doi.org/10.1080/00224545.1977.9713277

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, *36*(04), 393–414. https://doi.org/10.1017/S0140525X12000660

Sutskova, O., Senju, A., & Smith, T. J. (2022). Impact of Video-Mediated Online Social Presence and Observance on Cognitive Performance. *Technology, Mind, and Behavior*, *3*(2: Summer 2022). https://doi.org/10.1037/TMB0000023

Sutskova, Olga (2023). Cognitive impacts of social virtual reality: disentangling the virtual mere presence and audience effect. PhD thesis, Birkbeck, University of London. https://doi.org/10.18743/PUB.00051420

Uziel, L. (2007). Individual differences in the social facilitation effect: A review and meta-analysis. *Journal of Research in Personality*, *41*(3), 579–601. https://doi.org/10.1016/j.jrp.2006.06.008

Von Der Pütten, A. M., Krämer, N. C., Gratch, J., & Kang, S. H. (2010). "It doesn't matter what you are!" Explaining social effects of agents and avatars. *Computers in Human Behavior*, *26*(6), 1641–1650. https://doi.org/10.1016/J.CHB.2010.06.012

Wang, X., Lu, K., Runco, M. A., & Hao, N. (2018). Break the "wall" and become creative: Enacting embodied metaphors in virtual reality. *Consciousness and Cognition*, *62*, 102–109. https://doi.org/10.1016/J.CONCOG.2018.03.004

Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3), 225–240. https://doi.org/10.1162/105474698565686

Wolf, L. K., Bazargani, N., Kilford, E. J., Dumontheil, I., & Blakemore, S. J. (2015). The audience effect in adolescence depends on who's looking over your shoulder. *Journal of Adolescence*, *43*, 5–14. https://doi.org/10.1016/j.adolescence.2015.05.003

Zajonc, R. B. (1965). Social Facilitation. *Science*, *149*(3681), 269–274. https://doi.org/10.2307/1715944

Zanbaka, C. A., Ulinski, A. C., Goolkasian, P., & Hodges, L. F. (2007). Social responses to virtual humans. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '07*, 1561. https://doi.org/10.1145/1240624.1240861

Zender, R., Knoth, A. H., Fischer, M. H., & Lucke, U. (2019). Potentials of Virtual Reality as an Instrument for Research and Education. *I-Com*, *18*(1), 3–15. https://doi.org/10.1515/ICOM-2018-0042

Appendix A
Companion Animation Specifications

All the companion objects in **Figure 1A** (present humanoid companion **(a)**, present non-humanoid camera **(b)**, visually absent (none) non-social fan **(c)**) were animated in a Unity game engine replicating the movement of the object identically both in time and motion. To do so, the x, y, and z axis parameters of the humanoid companion's neck and head motion were replicated to the axis parameters motion of the camera on the tripod and the motion of the fan on its stand. Through this application, the companions' dynamic motion towards the participants and their performance screen, as well as a motion for non-engaging with participants and their performance were identical, performed by the different virtual objects. When the monitoring condition was "not monitoring", all objects' motion was identical by lowering its main component (humanoid head, non-humanoid camera on tripod, non-social fan head on its stand) down to face the floor 45 decrees. In the case of the humanoid companion, the 45 degrees was showing them looking into their virtual laptop, disengaging from the participant. All dynamic behaviours of the companion objects were piloted alongside colleagues assuring all objects' movements were not out of place or unrealistic.

The animation of the companion's behaviours (including non-social companion: fan) was set to loop after each monitoring trial. Each looping animation lasted 80 seconds, which was the maximum time a monitoring/not monitoring condition would last, based on the 20 trials 4 seconds each assigned by the experimental design. There were two animated looping behaviours, the watching down behaviour for not monitoring blocks, and the observing behaviours in monitoring blocks. The transitions between the blocks were animated as follows.

a)   **Present Humanoid Companion**   b) **Present Non-Humanoid Companion**

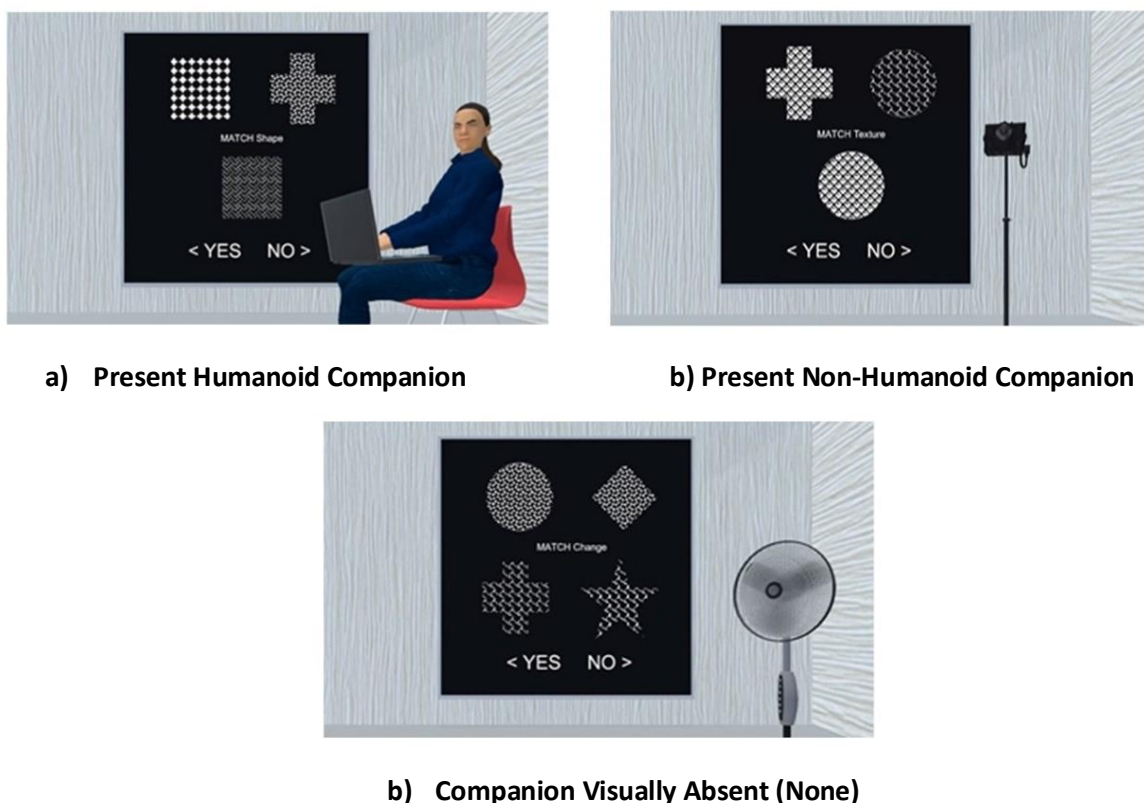b)   **Companion Visually Absent (None)**

Figure 1 A: The Companion Monitoring Presence Conditions, from the Participant's Perspective Within the Immersive Space.

When a Monitoring block ended and a Not Monitoring block began, the companion object would turn to the participant and then turn down, when the onscreen instruction said the next block condition. When a not monitoring block ended, and the monitoring block began, the virtual companion lifted their main component (head, camera, fan head) to an 85-degree angle turning towards the participant. The animation was programmed smoothly so as not to frighten the participant. The behaviours were overall animated to be smooth, natural as possible, and not purposefully distracting. Considering the participants had to focus on the task, not the object, the repetitive looping of animation seemed sufficient for animating social presence. The recurrence of looping behaviours was not picked up during piloting with peer researchers. When asked whether participants noticed any repetition, very few participants said they did. The only participant who did report some repetitive motion was the one who also suggested they did not pay attention to the task, but rather watched the companion. They were removed from the analysis.

Appendix B

The Summary of Means (M) And Standard Errors (SE) For All The Levels Of Experimental Design Per Each Factor.

Table B1: The Means (*M*) and Standard Errors (*SE*) of per cent accuracy (%) at all the levels by each factor.

Companion Agency	Companion Visual Presence	Monitoring	Difficulty	*M*	*SE*
Human	None	Monitored	Easy	95.30	1.19
			Difficult	88.75	2.64
		Not Monitored	Easy	95.30	1.26
			Difficult	89.15	2.75
	Present Non-Humanoid	Monitored	Easy	94.59	1.19
			Difficult	91.17	2.64
		Not Monitored	Easy	95.73	1.26
			Difficult	93.88	2.75
	Present Humanoid	Monitored	Easy	93.82	1.22
			Difficult	81.90	2.71
		Not Monitored	Easy	95.78	1.30
			Difficult	84.77	2.83
AI	None	Monitored	Easy	95.17	1.22
			Difficult	85.52	2.71
		Not Monitored	Easy	94.57	1.30
			Difficult	87.78	2.83
	Present Non-Humanoid	Monitored	Easy	95.03	1.22
			Difficult	88.69	2.71
		Not Monitored	Easy	94.72	1.30
			Difficult	90.05	2.83
	Present Humanoid	Monitored	Easy	88.91	1.26
			Difficult	81.09	2.80
		Not Monitored	Easy	91.35	1.34
			Difficult	77.91	2.92

Table B2: The Means (*M*) and Standard Errors *(SE)* of reaction times (RT, milliseconds) for all the levels by each factor.

Companions Agency	Companion Visual Presence	Monitoring	Difficulty	*M*	*SE*
Human	None	Monitored	Easy	1574.08	64.26
			Difficult	2077.51	70.04
		Not Monitored	Easy	1496.83	63.14
			Difficult	2017.70	72.33
	Present Non-Humanoid	Monitored	Easy	1590.04	64.26
			Difficult	2011.51	70.04
		Not Monitored	Easy	1562.18	63.14
			Difficult	1943.01	72.33
	Present Humanoid	Monitored	Easy	1762.29	66.12
			Difficult	2237.10	72.07
		Not Monitored	Easy	1702.37	64.97
			Difficult	2204.80	74.43
AI	None	Monitored	Easy	1697.76	66.12
			Difficult	2156.81	72.07
		Not Monitored	Easy	1678.91	64.97
			Difficult	2138.99	74.43
	Present Non-Humanoid	Monitored	Easy	1639.89	66.12
			Difficult	1940.44	72.07
		Not Monitored	Easy	1548.28	64.97
			Difficult	1900.82	74.43
	Present Humanoid	Monitored	Easy	1726.71	68.15
			Difficult	2077.25	74.29
		Not Monitored	Easy	1731.60	66.97
			Difficult	2052.64	76.72