

BIROn - Birkbeck Institutional Research Online

Sutskova, Olga and Senju, Atsushi and Smith, Tim J. (2022) Impact of video-mediated online social presence and observance on cognitive performance. *Technology, Mind, and Behavior* 3 (2), ISSN 2689-0208.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/48348/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>
contact lib-eprints@bbk.ac.uk.

or alternatively

SPECIAL COLLECTION: TECHNOLOGY IN A TIME OF SOCIAL DISTANCING

Impact of Video-Mediated Online Social Presence and Observance on Cognitive Performance

Olga Sutskova¹, Atsushi Senju², and Tim J. Smith¹

¹ Centre of Brain and Cognitive Development (CBCD), Birkbeck, University of London

² Research Center for Child Mental Development, Hamamatsu University School of Medicine



The onset of the coronavirus (COVID-19) pandemic has shifted most of the world toward remote working and education. As the world continues to embrace remote virtual communication in the post-COVID-19 era, it is crucial to investigate the impacts of online social copresence on cognitive performance. The present study investigated how the online videoconference presence of a virtual companion affects participants' performance on cognitive relational-reasoning tasks. The companion was either present and attentive to the participant, present but nonattentive, or absent. We manipulated the agency of a virtual companion, who was either a real human, an avatar controlled by a human, or an artificial intelligence (AI)-controlled agent. We hypothesized that the mere presence of a virtual companion, and the observance of participants' performance, would influence participants' performance. The results were broadly in line with our hypothesis that a mere presence of a virtual companion improved cognitive performance irrespective of their agency. However, the direction of the results did not support our prediction. We did not find a systematic impact of observance on cognitive performance, not supporting our second hypothesis. Participants performed best overall with an AI-controlled agent, next best with an avatar and worst with a real-human companion. We also observed that participants performed more accurately when a virtual companion was present but nonattentive, and faster when a virtual companion observed the participants, compared to when the participants performed alone. We conclude that online videoconference presence with a virtual companion, regardless of observance, temporarily enhances cognitive performance, and discuss the implications of these findings.

Keywords: cognitive performance, social facilitation effect (SFE), agent, avatar, online social video interaction

Supplemental materials: <https://doi.org/10.1037/tmb0000023.supp>

Special Collection Editors: C. Shawn Green, Nicholas David Bowman, and Tobias Greitemeyer.

Action Editor: Nick Bowman was the action editor for this article.

ORCID IDs: Olga Sutskova  <https://orcid.org/0000-0003-3222-524X>; Atsushi Senju  <https://orcid.org/0000-0002-8081-7170>.

Acknowledgements: We would like to thank our CBCD placement student (2020/2021) Billie Dale for lending us her time and testing all the 90 participants. Her diligence and attention to detail have contributed to the project significantly.

Funding: This research is part of the PhD thesis, which has been funded by a grant from the Economic and Social Research Council (ESRC), UCL, Bloomsbury and East London Doctoral Training Partnership (UBEL-DTP) funding body (1942074).




Disclosures: The authors certify that the research was conducted in the absence of any commercial or financial influences that could have led to perceived or potential conflicts of interest.

Registered Report (Approved in Principle): <https://osf.io/mbrzt>

Data Availability: The collated anonymized participant data will be publicly available on our OSF project page (<https://osf.io/d5ers>) on the date of publication. The analytic methods of the study, including minor deviations from the Registered Report, are publicly available and fully disclosed in the manuscript. The additional supplementary information

(testing pipeline) is publicly available on our OSF project page (<https://osf.io/d5ers>). The materials for the present study are fully disclosed and referenced in the manuscript. The edited template of the virtual character used in the manuscript and its technical specifications are available upon request to the main author.

Open Science Disclosures:

-  The data are available at <https://osf.io/d5ers>
-  The experiment materials are available at <https://osf.io/d5ers>
-  The preregistered design and analysis plan is accessible at <https://osf.io/mbrzt>

Open Access License: This article has been published under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. Copyright for this article is retained by the author(s). Author(s) grant(s) the American Psychological Association the exclusive right to publish the article and identify itself as the original publisher.

Contact Information: Correspondence concerning this article should be addressed to Olga Sutskova, Centre of Brain and Cognitive Development (CBCD), Birkbeck, University of London, Malet Street, London, WC1E 7HX, United Kingdom. Email: olga.sutskova.17@ucl.ac.uk

Since the coronavirus (COVID-19) pandemic onset and the subsequent lockdowns, many people have had to shift their habitual education, social and working environments into their homes, connecting to the rest of the world remotely through video chat and other virtual media. Adaptation to these sudden changes has led to novel digital interaction practices. Anecdotally, some people preferred to work alongside live videos of their colleagues working in the background, stating that the practice is motivating and keeps them more focused. An increase in videoconferencing has, however, also given a rise to reports of videoconference fatigue (i.e., Zoom fatigue), partially attributed to cognitive load exhaustion related social processing of online video communication (Bailenson, 2021). As society heads toward a more remotely interconnected world and the concerns over well-being increase, it is important to understand how remote social interactions affect our cognitive processing and behaviors.

Social interaction with others is important and affects activation in both the social and cognitive-executive regions in the brain (Jack et al., 2013). Empirical findings suggest that the presence of other people seems to magnify participants' subjective perception of their efforts on tasks, as well as increase the belief over their effort exertion to the teams' accomplishments (Steinmetz et al., 2016). Furthermore, research into a phenomenon known as the social facilitation effect (SFE; Zajonc, 1965) argues that the effects are not solely subjective, as both behavioral and cognitive performance is altered in the presence of other people (Bond & Titus, 1983; Claypoole & Szalma, 2018; Cottrell et al., 1968; Platania & Moran, 2001; Rajecki et al., 1977; Schmitt et al., 1986; Wolf et al., 2015). Therefore, the mere perception of being within a social context can change cognitive outcomes during the interaction.

Considering that there is an interplay between social processing and cognitive functioning, our current paper tests whether online videoconference-based social presence and monitoring by the virtual companions (human, human-controlled avatar, artificial intelligence (AI)-controlled agent) changes participants cognitive performance in real-time. We explore the impact of the phenomenon of the SFE on cognitive performance, with the focus on two processes proposed to generate the SFE, the mere presence effect (MPE), and the audience effect (AE).

Previous research into the SFE highlights two possible processes through which the social presence of other people alters human performance (Guerin, 1986; Guerin & Innes, 1984), eliciting the phenomenon called SFE. Firstly, the studies found that when participants believed someone is watching them perform (not co-performing), they tended to perform better on cognitive tasks in which they felt competent, and often worse on the tasks that they perceive as challenging or unknown (Bond, 1982). This process, which is hypothesized to be the response to social attentiveness, is called the audience effect (AE; Hamilton & Lind, 2016; Wolf et al., 2015). Secondly, studies indicate that the mere presence of someone else could itself be sufficient to change participants performance, worsening performance on challenging tasks, and facilitating the easier tasks. The process through which cognitive performance changes as a response to others presence, not-reliant on their attentive observation, is called the mere presence effect (MPE; Platania & Moran, 2001; Rajecki et al., 1977; Schmitt et al., 1986).

It seems that the mere belief about the social context a person might find themselves in dictates their response to it. The theory

of planned behavior suggests that people often adjust their behaviors based on their expectations of the context, what they believe is expected of them within that context, and by whom (Ajzen, 2011). SFE shows how these contextual adjustments might indeed benefit an individual on some tasks within a particular social context, yet might also have unexpected detrimental effects under another. The two processes presented above (AE and MPE) which may generate the phenomenon of SFE, although similar in performance outcomes, are possibly elicited through differential cognitive mechanisms driven by different social contexts. The AE is believed to be driven by increased social mentalizing (inferring another person's state of mind, e.g., what they think, know, or believe), fuelled by the reputation management strategies in the face of possible judgment from another person (Frith & Frith, 2007; Hamilton & Lind, 2016; Tennie et al., 2010). In contrast, the MPE does not seem to rely on higher-order social mentalizing (as it has been shown to also affect nonhuman animals, Zajonc, 1965), but rather on the attentional vigilance driven by the uncertainty of another conspecific action during the copresence in the same environment (Guerin, 1986). Indeed, nonhuman-primate studies show that when in presence of another conspecific, brain activation during task performance is significant in attentional, but not sociomotivational regions (Monfardini et al., 2016).

Undoubtedly, the beliefs about the social environment seem to dictate individuals' responses. But what aspect of human agency is important for a conspecific social response? Do people actually need another *human* in the environment, or can we simulate human presence, eliciting conspecific social processing such as AE or MPE?

With the emergence of artificial intelligence (AI) driven digital companions (agents) and human-controlled virtual characters (avatars), the line between computer and human social interaction becomes blurred. The anonymity between embodied (having visual interactive presence) avatars often facilitates more intimate disclosure than real-human face-to-face interaction (Green-Hamann et al., 2011), yet when participants feel identifiable to an anonymous observer (such as avatar), they tend to moderate their social responses (Joinson, 2001). The trust levels between avatars seem to be similar to trust for real humans, yet the brain seems to mentalize (engage in the inference of another person's mental states) less when interacting with avatars than real humans (Riedl et al., 2014). Interactive on-screen AI embodied conversation agents have been shown to reduce some levels of loneliness (Ring et al., 2015). However, inconsistencies in peoples' preferences and responses to virtual social AI companions show that AI-agents that are reliably perceived as meaningful social partners do not currently exist (Loveys et al., 2019).

Studies examining the social impact of virtual embodied agents and avatars on the cognitive performance of human participants have looked into the phenomenon of SFE with inconsistent results. Research of SFE within immersive virtual environments (IVE) predominantly reports improved cognitive performance when participants believe they are in the presence of humanoid avatars, but not agents (Hoyt et al., 2003; Okita et al., 2007). Other studies, however, report that agents' presence elicits SFE (Park & Catrambone, 2007; Zambaka et al., 2007), as long as the agent is humanoid (Garau et al., 2005) and displays human-like motion (Wellner et al., 2010). One possible explanation for the inconsistent findings is that virtual studies explore SFE as a generalized social impact effect, without separating or contrasting the engagement of

two processes that would elicit SFE, the MPE or AE. The consideration of these two processes could unravel the nuances of the social impact of different types of virtual assistants and help us to understand more about the intricacies of SFEs.

Critically, the distinction between a human-minded avatar and a not human-minded agent generates separate predictions for the engagement of AE and MPE, which are hypothesized to be driven by different cognitive mechanisms. Whilst both the embodied agent and avatar might share the same visual features (both in form and motion), the distinguishing feature of an avatar over an agent, is its ability to judge participants performance from the perspective of another person. Subjective reports of user experiences in virtual communication, support the notion that participants do not expect AI-intelligence to exhibit judgment similar to its human counterpart (Gratch et al., 2014; Pickard et al., 2016). As the AE is hypothesized to be (at least partially) subserved by mentalizing over others judgment during monitoring, we would expect AE to be engaged in eliciting SFE in an avatar (i.e., a counterpart with the capacity for social judgment), but not agent condition.

In contrast to AE, which requires an observer with the human ability to reflect and judge, the process of MPE is believed to be driven by attentional mechanisms aroused by the uncertainty of the actions of copresent conspecific. Although AI-agents are mostly not expected to exhibit judgment, embodied AI-agent interaction does seem to arouse social attention and (in some contexts) social-reward brain networks (Pfeiffer et al., 2014). Explanations for the social responses elicited by AI-agents comes from theories of human-AI communication, such as the media equation theory (Reeves & Nass, 1996) and the model of social influence (Blascovich, 2002). Both theories agree that social responses are mostly reflexive, applying human-communication heuristics to a human-like social situation, while consciously understanding that communication occurs with nonhuman “mind.” Currently, the embodied AI-agents can simulate human-like social behavior and motion reflecting their autonomous action. Considering that the mechanisms underlying MPE are hypothesized to involve social attention related arousal toward the copresence of other people, we predict that visual presence of autonomous AI-agents, just like the presence of human-driven avatars, might elicit MPE, in contrast to when performing alone.

Contrasting visually identical agents and avatars has the unique advantage of separating the human mind from a virtual body. Using this method, the present study is designed to test two hypothesized processes that elicit the phenomenon of SFE (i.e., MPE and AE), by testing the impacts of virtual social copresence on cognitive performance and exploiting the unique societal context which led to the recent boost of video-mediated online social interaction. However, both virtual character conditions are assumed to have a lesser social impact than the video-based live presence of another human, according to the threshold model of social influence (Blascovich, 2002). This model suggests that due to the high levels of social realism and personal relevance experienced by the participant during a live videoconference, real-human video communication will always have a social advantage over a virtual character’s visual presence.

The present study will look into the levels of social influence threshold, contrasting how the real-time video-based presence of the agent, avatar, and a real-human (live video) impact participants’

cognitive performance. Participants will be exposed to either of the three types of social companions, as they perform a cognitive task through a currently widely used online messenger software (Zoom.us) from the comfort of their homes. Most critically, the study is not designed to directly test the mechanisms underlying the processes of MPE and AE. Instead, the present study is designed to systematically manipulate social context, informed by the hypothesized cognitive mechanisms underlying MPE and AE. By systematically contrasting cognitive performances between the social contexts which should selectively engage MPE and AE process with those which should not engage these processes, we tested whether working alongside human, avatar or agent companions online, elicited the cognitive performance effects characteristic to SFE.

In accordance with the canonical manifestation of SFE, we expect that the SFE’s will manifest as better performance on easy tasks (lower RT, higher accuracy) and worse performance on difficult tasks (higher RT, lower accuracy) when accompanied by a conspecific in contrast to being alone. The significant difference between the easy and difficult relational reasoning paradigm (RRP) conditions will be analyzed as part of the omnibus analysis (difficulty).

Hypothesis 1: Mere presence effect. The MPE hypothesizes that SFE rises due to uncertainty of conspecific actions within the shared environment, irrespective of who they are or whether they are actively judging or observing the participant. We, therefore, predict that the MPE-related SFE will occur only when the Conspecific is present (vs. no observer present), regardless of whether they attend to the participants (nonattentive or attentive). In our manipulation, we expect MPE in all three conspecific conditions (real-human, avatar, agent). Please see the Analysis Plan (Appendix A) for interactions of interest.

Hypothesis 2: Audience effect. The AE assumes that SFE arises due to mentalizing processes relating to others judgment of one’s performance and that these mentalizing processes only occur when the participant believes the attentive observer is capable of mentalizing, that is, real-human or avatar but not agent. Therefore, we predicted that we will only observe the SFE (AE) in the presence of an attentive observer (attentive vs. nonattentive and no observer) with the capacity to mentalize, that is, real-human or avatar but not agent. Please see the Analysis Plan (Appendix A) for interactions of interest.

For exploratory analyses on social impact, see Analysis Plan (Appendix A).

Method

Participants

Out of 90 participants tested, data from 54 adult participants (18 per between-subjects group), 44 female (10 male), mean age $M = 26.94$ ($SD = 5.87$), age range 19–41 years were entered into the final analysis (see reasons for participant exclusion below). Data were gathered from an opportunistic sample of university students and employed adults, self-reported as neurotypical with no clinical diagnosis in autism spectrum disorder (ASD) and social anxiety. The target sample size ($N = 54$) was estimated using G*Power, at

$1 - \beta = .8$, $\alpha = .05$, Cohen $f = .44$. The effect size was estimated from significant (Task \times Audience) AE-based interaction in Dumontheil et al. (2016). Participants who did not understand the cognitive task or did not believe the conspecific implication were excluded from the analysis (see “Participants Exclusion” below for more information). Access to a personal computer, stable internet connection, and a good lighting source at participants’ homes was a formal requirement.

We had to test 90 participants instead of the preplanned recruitment target of 70, due to unexpected difficulties encountered during the COVID-19 pandemic (see Appendix B: Deviations From Registered Report for more details). Out of 90 participants who attended the study, 31 were excluded due to the issues related to remote in-home testing, such as technical issues (internet, camera, and computer problems) and home-based distraction (alarms, doorbells, street noise, other household residences). Out of the remaining 59 participants, five participants were removed for not believing the conspecific manipulation (see “Participants Exclusion” below).

Design

As per our Registered Report (Sutskova et al., 2020), we have systematically manipulated three independent factors, the identity of conspecific (between-s; real-human, avatar, agent), the degree of observance (within-s; no observer, nonattentive, attentive), and task difficulty (within-s; easy, difficult). Cognitive performance was measured in reaction times (RT, for accurate responses only) and percent accuracy (%) during the performance of RRP (Dumontheil et al., 2016). To reduce the possible reflexive-random responses, both per cent of accurate responses and RT were only counted for the trials which correct answer RTs are within the 99.7 percentile (3 SD , per each difficulty condition), and over 250 ms from the stimuli onset, based on the average RT for keyboard response to visual stimulus onset (Jain et al., 2015). The original estimation of 95 percentile was overwritten due to difficult testing conditions (see Appendix B: Deviations From Registered Report).

Participants Exclusion

After the study completion and before debriefing, all participants were asked about their subjective experience of the virtual interaction. To make sure the participants believed their conspecific

condition (real-human, avatar, agent) the researcher asked the participant directly whether they believed the conspecific was either AI- or researcher-driven. The belief response was noted down as binary YES (1)/NO (0). Participants then reported thoughts about their subjective experiences under the social presence conditions (attentive, nonattentive). The subjective reports were noted down alongside the participant number and the binary belief response. Only participants who believed their interaction occurred with the assigned conspecific, with conspecific performing an assigned action (nonattentive, attentive) were included in the final analysis. Uncertain or disbelieving participants were excluded.

Software and Stimuli

Zoom video chat messenger (Zoom.us) was used so the researcher and participant could communicate remotely. Screen-share was used to visually project the participant’s view of the task to the researcher. An online experimental task engine (Gorilla.sc) was used to enable participants access to the experimental RRP task at home.

For the real-human observer, the researcher used their live video feed. For the virtual observers, the same character was used both for Avatar and Agent (Figure 1). The virtual character in the study was a visually modified free template illustration (Cassandra) provided by the Adobe Character Animator (www.okaysamurai.com/puppets). The character was controlled via Adobe Character Animator software in real-time. The software tracked the researcher’s live motion and gaze-shifts through their web camera (both for agent and avatar), as well as lip-synched to the researcher’s speech (only for avatar condition). The character could look at the participant and their performance or look away. Active eyes were used in the attentive condition when noting down participants performance. The character was turned away when not attending to the participant’s performance (nonattentive).

The cognitive task used for the experiment was RRP, a visual pattern matching logic task, adapted from Dumontheil et al. (2016), proven to elicit SFE. The task consisted of two difficulty levels of cognitive load, easy and difficult (visual examples in Figure 3, p. 6). During the easy condition, participants saw three images and were required to match either the shape or texture of the top two images to the image on the bottom. In the difficult condition, participants were exposed to four images and were asked to match whether the way the top images changed in the same way as the bottom two (either shape or texture). There were 40 easy and difficult randomized trials

Figure 1

Virtual Character (Avatar or Agent) During Live Marking (Attentive) With Open Active Eyes or Turned Away During No Marking (Nonattentive)



per each of the three condition blocks, 120 trials in total. Participants used left and right keyboard presses to respond to the task presented for a maximum of 3.5 s per trial.

Participants' personal computers (laptops) needed to be equipped with a working web camera and a keyboard to test participants from their own homes. The experiment was screen-captured.

Conditions and Social Deception

To avoid any confounds based on participants' guessing the experimental manipulation, minimal social deception was made (ethical approval granted). The participants were told a cover story that we were piloting a new online-testing-software that can track participants performance based on messenger screen-share data and predict whether their responses were going to be correct or not based on their webcam-based eye-gaze. Participants were told that they would be informed when marking (pen icon: attentive observer) and no marking sessions occur (crossover pen icon: nonattentive observer) via on-screen notification icons. Please see the on-screen instructions and the study conditions pipeline on our OSF project page: <https://osf.io/d5ers>.

Observance Conditions (*Within-Subjects*)

Attentive, nonattentive, no observer. Conditions were counter-balanced, with no observer being either first or last condition, and nonattentive and attentive conditions randomized.

For the attentive condition, participants were told that their virtually copresent partner would be manually marking their performance for future quality control of the automated online-testing-software. In the nonattentive condition, participants were told that their partner would be visually present but not attending to their task. In the no observer condition, participants were told the researcher needed to set up software; therefore, all media sharing needed to be switched off (no camera or screen-share). Participants performed the task alone and told the researcher through the microphone when they finish.

Conspecific Conditions (*Between-Subjects*)

Participants were randomly assigned to real-human, avatar, or agent groups.

Real-human observer group performed next to a live video of the researcher marking down participants performance through screen-share during attentive condition. The researcher was looking away from the participant during nonattentive conditions, busy working on other tasks.

The avatar was controlled by the researcher to pay attention to participants answers during the attentive condition, providing natural character motion and eye gaze shifts based on researchers' natural gaze. During the nonattentive condition, the motion of the character was present and based on the researchers' natural motion. However, in contrast to the attentive condition, the avatar was turned away from the participant showing the participant its back (see Figure 1). Participants were told that when the researcher does not face their laptop camera, the avatar is programmed to turn around, avoiding participants belief that observation occurs during nonattentive conditions.

In the agent condition, the character was controlled by the researcher in the same way as for the avatar condition. Participants

were instructed that the agent was controlled by an in-house AI algorithm designed to mark user performances remotely using live on-screen data. Participants were told that at one point, the algorithm will access their screen-share and video data to make predictions about their performance (attentive), and other times the program will just run in the background without analyzing their performance in real-time (nonattentive). Participants were told that the AI-agent was preprogrammed to show active-gaze reflecting what they read from screen-share and video data. Therefore, participants assumed that when the AI live data marking occurs (attentive) the agent will be actively observing, and when the data marking does not occur (nonattentive) the agent turns away.

Emergency communication was always kept through the audio channel (no lip sync for agent condition). The instructions throughout the experiment were delivered by onscreen text instructions for all conditions.

Procedure

All participants logged in to the browser-based experimental software (Gorilla.sc) and followed task-screen and messenger preparation instructions. Please see our OSF page infographic for the illustration of the study sequence pipeline: <https://osf.io/d5ers>. During setup, the participant activated Zoom.sc messenger screen-share, so the researcher would see their task screen. All the participants had to position the researcher's messenger video window within a designated region on the right side of the screen (Figure 2). Participants were told that the video screen positioning is a requirement for consistency between participants experimental layouts. The researcher highlighted that this layout requirement made sure that the researcher's video screen would not interfere with the main task region and stimuli.

After the experimental setup confirmation from the participant, the researcher introduced the cognitive task RRP, at both easy and difficult levels, with a short practice (5 trials each, ensuring task understanding) and a questions session. Practice session followed with gaze calibration and then the experimental task, starting either with no observer condition (see Figure 3 for participant view of the condition) or either attentive or nonattentive conditions in which there was copresence with the virtual conspecific.

Data Availability Statement

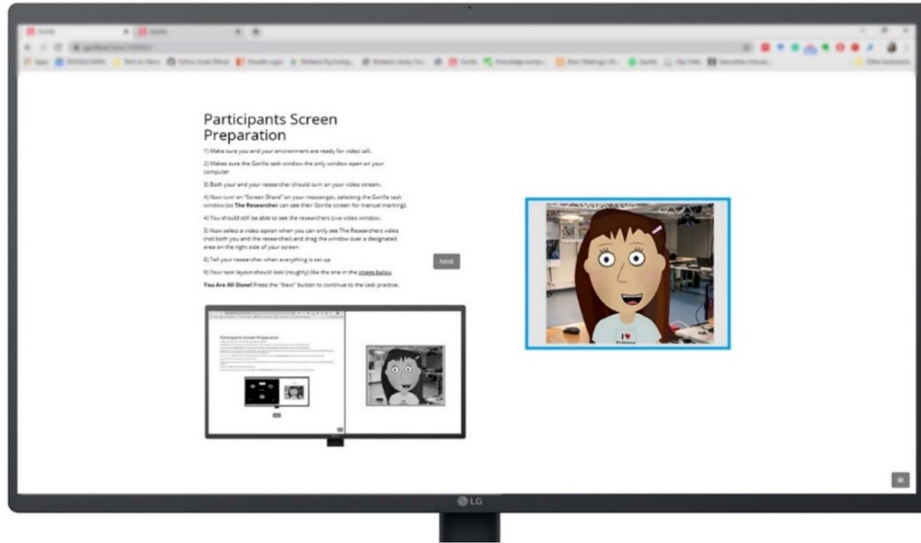
The collated anonymized participant data will be publicly available on our OSF project page (<https://osf.io/d5ers>) on the date of publication. The analytic methods of the study, including minor deviations from the Registered Report, are publicly available and fully disclosed in the manuscript. The additional supplementary information (testing pipeline) is publicly available on our OSF project page (<https://osf.io/d5ers>). The materials for the present study are fully disclosed and referenced in the manuscript. The edited template of the virtual character used in the manuscript and its technical specifications are available upon request to the main author.

Results

The analyses conducted in this section follow a preregistered Analysis Plan (Appendix A), part of the Registered Report (Sutskova et al., 2020). For slight deviations from the Analysis Plan, see Deviation From Registered Report (Appendix B).

Figure 2

Participants View of the Preparation Screen, With Observers' Video Fitted on the Right



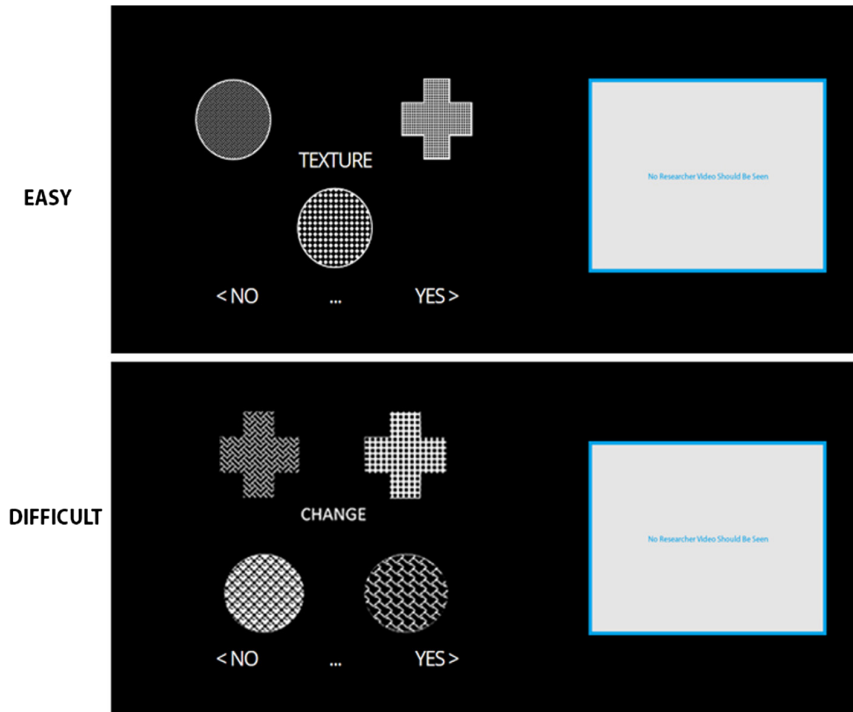
Participants with an accuracy less than 3 *SD* (99.7%) below the sample mean (originally 2 *SD*, see Appendix B), and under 250 ms in average RT, were excluded from the analysis. Only participants who firmly confirmed belief in the manipulation (see Participants Exclusion in Methods) were included in the final analysis. As per

Analysis Plan, the statistical analyses were run on $N = 54$ participants, with 18 participants per conspecific group.

Sphericity of data was confirmed ($p > .05$) between and within the conditions of interest both for accuracy and RT. Levene's test showed a slight homogeneity deviation in accuracy between the

Figure 3

Participants Task View of the No Observer (Easy, Difficult) Condition, With No View of the Researcher in the Window on the Right



Note. For attentive and nonattentive conditions, the conspecific companions live video feed will be present.

conspicuous conditions for nonattentive difficulty, $p = .023$, for RT homogeneity confirmed ($p > .05$).

Two mixed three-way analysis of variances (ANOVAs), 3 conspecific \times 3 observance \times 2 difficulty, were conducted on the accuracy (%) and RT (ms) separately, followed up by planned analyses (see Analysis Plan: Appendix A). The main body of the Results section focuses on reporting the outcomes of the hypothesized effects (MPE, AE, and conspecific context effects) predicted in the preregistered Analysis Plan (Appendix A) and any additional significant results that were not originally predicted. For the remaining nonsignificant findings see Appendix C, for figures of overall mean (M) and 1 standard deviation (SD) for the three-way ANOVAs see Appendix D.

Difficulty

Before testing the hypotheses presented above, we investigated the main effects of difficulty to confirm that RRP task difficulty levels employ different levels of cognitive effort (i.e., difficult tasks were indeed more “difficult” than easy tasks), which is prerequisite of the RRP. A three-way ANOVA indeed indicated a significant main effect of difficulty, both for accuracy and RT in accordance with expectations of the RRP task, with expected direction. The performance on the easy condition was more accurate and faster than those on the difficult condition: Accuracy, $F(1, 51) = 48.85$, $p < .001$, $\eta_p^2 = 0.49$ (Easy: $M = 92.2$, $SD = 6.200$, Difficult: $M = 80.60$, $SD = 14.0$), and RT $F(1, 51) = 148.66$, $p < .001$, $\eta_p^2 = 0.75$ (Easy: $M = 1396.25$, $SD = 291.37$, Difficult: $M = 1907.7$, $SD = 422.87$).

Mere Presence Effects

For the MPE hypothesis, we predicted that the three-way ANOVA analysis would indicate a significant two-way observance \times difficulty interaction. In percent accuracy, there was a significant observance \times difficulty interaction as predicted, $F(2, 51) = 3.181$, $p = .047$, $\eta_p^2 = .059$. Planned follow up comparisons, between the no observer and observer present conditions (the nonattentive and attentive observer conditions combined), revealed that the performance accuracy for the difficult task was significantly higher in the observer present conditions ($M = 82.04$, $SD = 14.82$) in contrast to no observer conditions ($M = 77.60$, $SD = 15.35$), $t(1, 53) = 2.85$, $p = .006$. Note that the direction of the effect was opposite to our Hypothesis 1, which predicted that the performance in the difficult condition would decrease (and would increase in easy condition) in the presence of an observer. There were no significant differences for the easy conditions $t(1, 53) = 0.051$, $p = .96$, between the observer present conditions ($M = 92.18$, $SD = 6.0$) and no observer ($M = 92.13$, $SD = 8.45$) conditions.

For RT, there was no significant MPE related observance \times difficulty interaction, in contrary to prediction, $F(2, 51) = 2.31$, $p = .104$, $\eta_p^2 = 0.043$. The planned follow up comparisons (see Analysis Plan) between the no observer condition versus observer present conditions, revealed that similarly to accuracy, the RT for the difficult task indicated significantly better (faster) performance, $t(1, 53) = 2.35$, $p = .023$, during the online observer present ($M = 1876.08$, $SD = 425.93$) in contrast to no observer condition ($M = 1970.94$, $SD = 481.94$). Again, the direction of results was opposite our Hypothesis 1 prediction, which predicted that

performance in the difficult condition would decrease (and would increase in easy condition) in the observer present (MPE) condition. For the easy conditions, there were no significant difference between observer present ($M = 1387.60$, $SD = 309.10$) and no observer conditions ($M = 1413.55$, $SD = 306.96$), $t(1, 53) = 0.89$, $p = .38$.

Observance \times difficulty interaction in both accuracy and RT indicate significant performance improvement on the difficult, but not easy tasks, when performing alongside an online conspecific in contrast to performance alone (Figure 4). Although it is in line with H1 that mere presence of others influences task performance, the results did not support the directional prediction derived from the canonical SFE literature, that during observer present conditions, the performance will decrease in difficult condition and increase in easy condition, when compared to the performance in no observer conditions.

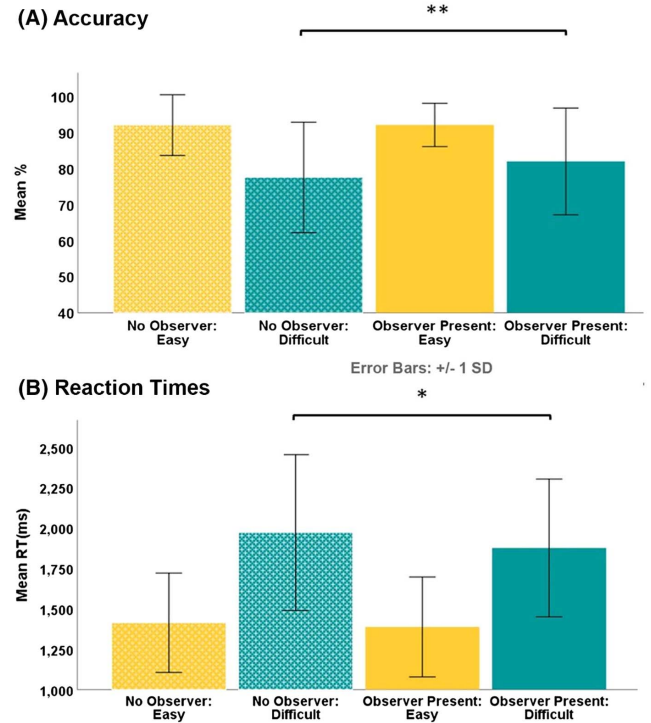
Audience Effect

For the AE related hypotheses, we predicted a significant three-way difficulty \times observance \times conspecific interaction, with a set of observance \times difficulty planned contrasts to be conducted within each conspecific group (see Appendix A: Analysis Plan).

There was no significant difficulty \times observance \times conspecific interaction, neither for accuracy $F(4, 53) = 0.572$, $p = .68$, $\eta_p^2 = 0.022$, nor for RT $F(4, 51) = 0.97$, $p = .43$, $\eta_p^2 = 0.037$, which did not support our hypothesis.

Figure 4

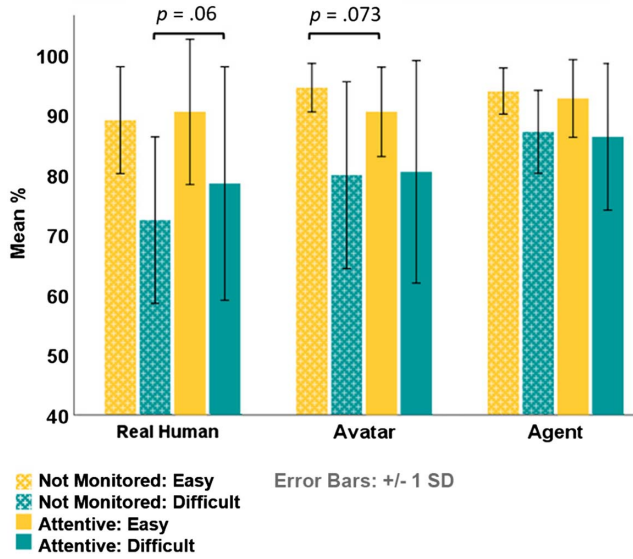
Mere Presence Effect Planned Contrast Descriptive Statistics (M , 1 SD) for Accuracy (A) and Reaction Times (RT; B)



Note. There is a significant increase in difficult performance in observer present conditions in contrast to no observer. The increase is present in both accuracy (A) and RT (B), * $p < .05$, ** $p < .01$, respectively. No social facilitation effect idiosyncratic interaction is present.

Figure 5

Audience Effect Planned Contrast Descriptive Statistics (M , 1 SD) for Accuracy



Note. Marginal significant change in the performance in real-human and avatar observer groups. The change however does not follow the social facilitation effect idiosyncratic interaction.

A planned AE comparison was performed within each conspecific group, comparing attentive observer (when the observer is marking) versus not monitored (no observer and nonattentive observer combined) conditions. We predicted that the AE (i.e., worse accuracy in difficult conditions, and better accuracy in easy conditions, in the attentive observer condition compared to conditions in which participants were not monitored), will emerge only in real-human and avatar conditions, but not in agent condition. For performance accuracy there was no significant observance \times difficulty interaction within either of the conspecific groups, real-human, $F(1, 17) = 1.67$, $p = .21$, $\eta_p^2 = 0.09$, avatar, $F(1, 17) = 2.96$, $p = .10$, $\eta_p^2 = .15$, or agent $F(1, 17) = 0.025$, $p = .88$, $\eta_p^2 < 0.001$, which did not support our hypothesis.

Planned comparisons for accuracy indicated that participants in real-human and avatar conditions showed marginal performance change between attentive observer and not monitored conditions (Figure 5). The real-human attentive observance marginally increased participants performance only in difficult condition ($M = 78.61$, $SD = 19.46$) in contrast to not monitored difficult condition ($M = 72.50$, $SD = 13.88$), $t(1, 17) = 2.024$, $p = .06$. In avatar condition attentive observance decreased performance only on easy condition, with easy condition being marginally worse under attentive observer ($M = 90.56$, $SD = 7.45$) than when not monitored ($M = 94.58$, $SD = 4.04$), $t(1, 17) = 1.91$, $p = .073$. However, none of these effects reached significance, and directions of nonsignificant effect were not consistent with each other.

As predicted, there was no significant change in agent groups (see Appendix E for the breakdown of AE t -tests).

For performance RT, there were no significant observance \times difficulty interaction within either of the conspecific groups, real-human, $F(1, 17) < .001$, $p > .99$, $\eta_p^2 < 0.001$, avatar, $F(1, 17) = 2.88$, $p = .11$, $\eta_p^2 = .145$, or agent $F(1, 17) = 0.58$, $p = .46$, $\eta_p^2 < 0.033$.

Planned contrasts indicated that only avatar group showed significant increase on the easy condition between attentive observer ($M = 1269.01$, $SD = 299.76$) and not attended ($M = 1393.57$, $SD = 320.62$) conditions, $t(1, 17) = 2.68$, $p = .016$ (see Figure 6), there were no significant difference in real-human and agent groups (see Appendix E for breakdown of AE t -tests).

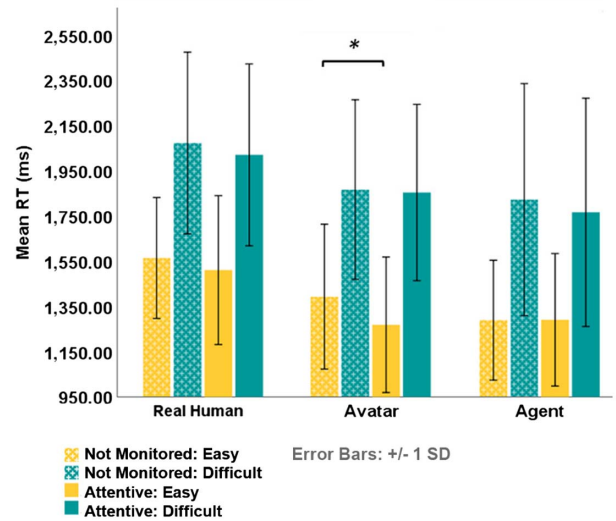
Overall H2 was not supported from our results, with no predicted three-way interaction between difficulty \times observance \times conspecific or observance \times difficulty interaction in real-human or avatar condition. However, there was weak support for a broader AE effect, that real-human and avatar, but not agent, conditions show a trend that performance in attentive condition is different from when participants were not monitored (no observer and nonattentive combined). Note that these effects only reached significance in one condition in RT and approached significance in two conditions in accuracy. These directions of results were however not in line with our directional prediction derived from the canonical SFE literature, which would predict worse performance in difficult trials and better performance in easy trials during attentive observer, versus when not monitored by the observer (AE).

Conspecific Context

A three-way ANOVA revealed a significant main effect of conspecific group, both in accuracy $F(2, 51) = 4.40$, $p = .017$, $\eta_p^2 = 0.147$ and in RT $F(2, 51) = 3.14$, $p = .052$, $\eta_p^2 = 0.11$. There was a linear increase in performance with decreasing “humanness” (a linear trend for accuracy, $p = .005$; and RT, $p = .02$), between the conspecific groups, with real-human groups performing overall

Figure 6

Audience Effect Planned Contrast Descriptive Statistics (M , 1 SD) for Reaction Times (RT)



Note. Significantly faster performance in avatar observer group only. No other conspecific groups performance change reached statistical significance. The significant change observed did not follow the social facilitation effect idiosyncratic interaction.

* $p < .05$.

worst, followed by avatar, and the best performance is in the agent group.

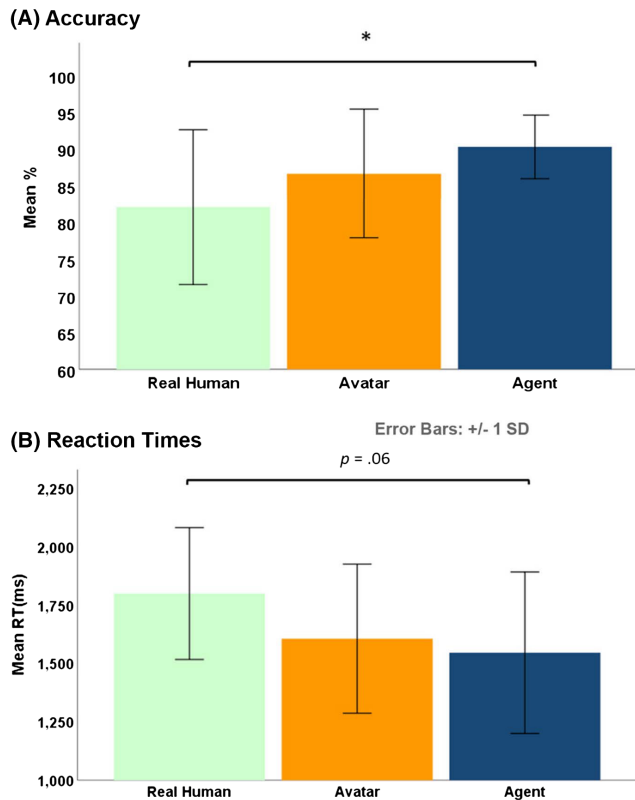
The proposed follow up contrasts (3 pairwise, Bonferroni corrected, see Analysis Plan in Appendix A) between the three conspecific groups in accuracy, indicated a significant difference between real-human and agent groups, $p = 0.014$, with no significant difference between agent and avatar ($p = .61$), or avatar and real-human groups ($p = .3$), see Figure 7A. The RT indicated a similar trend, with a marginal but nonsignificant difference between real-human and agent ($p = .06$), and no significant differences between avatar and agent $p > .99$, or real-human and avatar groups ($p = .22$), see Figure 7B.

Observance

A three-way ANOVA revealed a main effect of observance, for accuracy $F(2, 51) = 3.39$, $p = .037$, $\eta_p^2 = .062$, and a marginal effect for RT $F(2, 51) = 2.47$, $p = .090$, $\eta_p^2 = .046$.

Posthoc exploratory comparisons (Bonferroni corrected, 3 pairwise test) of a significant observance effect in accuracy indicated a significant quadratic ($p = .038$), but not a linear trend ($p = .128$), of

Figure 7
Effect of Conspecific Type (M , 1 SD) on Accuracy (A) and Reaction Times (RT ; B)



Note. A gradual improvement in performance as social influence decreases, (A) accuracy, (B) RT.

* $p < .05$.

performance change with the increase in attentive presence, in direction from no observer to nonattentive then to attentive observer (see Appendix D for breakdown of conditions). Nonattentive presence ($M = 87.64$, $SD = 8.84$) has significantly improved accuracy of performance versus no observer condition ($M = 84.87$, $SD = 10.2$), $p = .015$, with no significant difference between no observer and attentive observer ($M = 86.57$, $SD = 10.77$), $p = .128$, and nonattentive observer and attentive ($p = .36$).

Posthoc exploratory contrasts (Bonferroni corrected, 3 pairwise test) for marginal main effect in RT revealed a significant linear ($p = .019$), but not quadratic ($p = .74$) trend of performance change with the increase in attentive presence, with the slowest performance in no observer ($M = 1692.24$, $SD = 352.23$), followed by nonattentive ($M = 1644.84$, $SD = 382.37$), and fastest performance in attentive observer conditions ($M = 1618.85$, $SD = 340.02$). There was a significant performance difference between no observer and attentive observer, $p = .056$, and no significant difference between no observer and nonattentive ($p = .56$), and nonattentive and attentive observer ($p > .99$).

Discussion

The current preregistered experiment tested whether the phenomenon called social facilitation effect (SFE), which is often reported for face-to-face scenarios, also impacts cognitive performance during the increasingly common scenario of an online video meeting. Participants were asked to perform a quick-response visual logical reasoning task (RRP; Dumontheil et al., 2016), under different levels of confederate presence during an online video meeting. We compared how the perceived social agency of the online other (conspecific) impacted participant performance at different levels of social presence and attentiveness. The social impact was predicted using the threshold model of social influence (Blascovich, 2002). Participants had an online video interaction with one of the three different levels of human presence: highest being in a call with a confederate (real-human, realistic visual human presence), the middle being in call with a visually less realistic human-controlled animated avatar (implied human presence), and lowest being in a call with an AI-algorithm controlled animated agent (nonhuman presence). The social impact was tested based on the predictions derived from the theories on two processes eliciting the phenomenon of SFE, MPE (Rajcecki et al., 1977), and AE (Wolf et al., 2015).

Our results showed that during an online video meeting, the mere presence of another conspecific significantly altered performance on more difficult tasks, with performance becoming more accurate and quicker. The performance was not significantly affected by whether participants believed their performance was observed nor by the type of conspecific present. It was affected by the shared mutual video presence. Although our findings support the prediction of performance change in presence of a virtual online companion versus performing alone, the results did not support the predicted direction of our first hypothesis: SFE would manifest as an increase in performance for easy and a decrease in performance for difficult trials.

For the AE, we hypothesized that the participants' belief in being attentively observed would change their cognitive performance according to the SFE. Unlike the MPE, which we postulated was subserved by more primitive cognitive mechanisms, AE was

hypothesized to be subserved by mentalizing. Therefore, we predicted the engagement of AE on the manifestation of SFE in the presence of a human-mind (real-human and avatar), but not in the presence of nonhuman conspecifics (i.e., AI) companion (agent). Our results were numerically and broadly in line with our prediction that participant performance would change during attentive observation by the human-mind companions, but not in the agent group. However, the finding did not reach statistical significance, showed an inconsistent effect, and was not in the predicted direction from the canonical SFE literature. Hence, we cannot draw any firm conclusion on the hypotheses derived from AE, and more research needs to be conducted to explore the intricacies of possible differences in their social impact.

Whilst SFE is widely replicated, the classic direction of the effect (i.e., improved performance for easy trials and impaired for difficult) is not, with some studies showing directional effects similar to ours, rather than the canonical interaction (Hoyt et al., 2003; Okita et al., 2007; Zambaka et al., 2007). Meta-analyses support the notion that not all studies find these particular interaction effects or that effects on easy trials are statistically weak (Bond & Titus, 1983). The effects on difficult conditions might depend on how it was evaluated by the participants: whether the task is considered a challenge and therefore motivated by social others, or uncomfortably difficult and therefore elicits threatful arousal in presence of others. Depending on the task's level of difficulty, the performance might either increase or decrease in social presence, respectively (Blascovich et al., 1999). Therefore, how different levels of difficulty are perceived might impact performance and should be inspected further.

Alongside our main effects of interests, additional evidence on the social impact of different types of conspecifics can be derived from our planned exploratory analyses between the conspecific groups (no direction predicted in "Analysis Plan"). Our results showed a gradual overall decrease in accuracy and speed of performance as the level of "humanness" of conspecifics increased. Real-human groups performed worst, followed by avatar, and with the best performance in agent groups. A similar social influence linear (gradient) trend was reported in the study looking into effects of social support contrasting real-human, avatar, and agent, reporting most beneficial social support impacts in human, then avatar, and least in agent condition (Kothgassner et al., 2019). These social (conspecific) context effects, irrespective of observance (presence) type, is in line with the theory of planned behavior, which suggests that people adjust their behavior and expectations depending on their beliefs about the context they are planning to engage with (Ajzen, 2011). In our experiment, the instructions at the beginning of the study stated the social interaction context, under which the participant should anticipate performing (either real-human, avatar, or agent). Considering that participants did not know at which stage the observation will happen (randomized blocks) there is also a possible evaluation-anticipation effect relating to the conspecific type. In our study, the evaluation-anticipation in a more socially influential (real-human) context may have prompted more anticipatory evaluation-stress than the AI agent mediated (low social impact) condition. Indeed, some findings suggest that the cortisol-response peaks both during anticipatory and reactive social evaluation stresses (Engert et al., 2013), especially during upcoming social cognitive (mental)

evaluation (Dickerson & Kemeny, 2004). In the real-world settings, the higher levels of cognitive evaluations stress have been reported to have detrimental effects (at least acutely) on both cognitive performance (working memory: Angelidis et al., 2019; IQ-test battery: Elliot et al., 2011) and academic test performance (SAT scores: Cassady & Johnson, 2002). Our study shows that these impacts possibly generalize to online video social context and explain why the real-human condition was affected the most.

In addition to our main predicted effects, the results also showed the main effect of observance (presence) type overall. In contrast to performance when alone (no observer), participants performed significantly faster when attentively observed (linear trend as observance level increases), but more accurately when the virtual companion was video-present but not attending to their performance (quadratic trend). This is an interesting finding in itself, as accuracy and speed of performance might vary based on the type of presence. The knowledge that performance is being marked (in the attentive condition) might push participants to show off better performance (increasing the speed of the task), which could have led to a speed-accuracy trade-off, leading to the observed quadratic trend in performance accuracy. On the other hand, additional arousal of presence, without the requirement of performative action, might create relevant arousal for attention without performative distraction. Therefore, the level of perceived observance itself, regardless of the perceived level of task difficulties, could change the task performance of the participants. As this finding is exploratory, our speculations would merit further research.

Considering our findings, we can conclude that sharing an online video call with others impacts cognitive performance on a co-occurring task. At least for a limited duration (4 min per condition block), the participants' performance was enhanced when they performed more difficult tasks in the mere presence of an online conspecific, irrespective of their belief of whether the task is being attended by their online video companion or not. Additionally, our results were not inconsistent with the claim that when participants believe their performance is observed (attended to), the social influence of human-minded video companions (real-human, avatar) but not by AI-driven character (Agent—not human-minded, but visually identical to avatar) impact participants performance differently. However, as the differences were not significant according to our AE predictions, both the real-human and avatar impact differences and why these differences occur requires further investigation. For example, future exploration could investigate possible causal impacts of different levels of observed emotional reactivity, ambiguity, and anonymity levels of the communication partner, both in video and avatar forms.

It is important to highlight that the current article only aimed to dissociate MPE and AE by manipulating the social context during a task and not the underlying mechanisms giving rise to the two effects. Future work should directly target these hypothesized underlying cognitive mechanisms, such as reputation management strategies or attentional vigilance. However, a more robust systematic approach is still required to establish the direct causal relationships between these cognitive mechanisms and the two processes we focused on, which are hypothesized to engage in the manifestation of SFE.

Current findings highlight the importance of considering social factors when discussing cognitive performance and cognitive load during communal virtual and online video interaction. Although the present study investigated only the short-term effects of social-to-cognitive impacts, long-term effects should be investigated in-depth (see Zoom Fatigue), especially considering that a substantial amount of education and office work might shift toward a remote work-at-home approach.

It is worth reiterating that the agent companions were considered impactful enough to create a social presence-related performance boost, particularly the MPE, whilst maintaining the overall best performance of three conspecifics. Therefore, at least for short-term vigilance and performance boost, without additional social strain, social agents might be sufficiently engaging educational and work companions. Indeed, agents used for educational purposes have already proven to make the process more enjoyable and less stressful (Jin, 2010).

It is worth noting that the present study was mainly targeted at the educational and work sector where real-time high-intensity cognitive performance matters, therefore implications from our study should not be automatically extended to other sectors, particularly the health sector. As interest in remote work and services increases, so do the propositions for remote AI assistance in human-compassion-based sectors, such as healthcare support and therapy. As such AI support systems evolve it is important to consider the context in which virtual socialization occurs. Therefore, when researchers and practitioners discuss the possible benefits of AI agents in emotionally supportive settings (Fiske et al., 2019), they should also reflect on the evidence of poor educational and therapy retention outcomes during the COVID-19 pandemic, explained by the lack of in-person connection (Aboujaoude et al., 2021). Our study indicated that the real-human interaction, even when communication is remote, might still be the most impactful way to ensure greater remote social effect within a general context of monitoring and influencing cognitive performance. Future research should extend the current finding in sectors in which authentic human compassion and support are required for health and emotional outcomes (e.g., Kothgassner et al., 2019).

In conclusion, the present study demonstrated a social presence impact on cognitive performance during online video interaction, and that the levels of observed presence and perceived “human-ness,” such as the visual or implied presence or implied nonhuman presence, significantly vary cognitive performance outcomes.

References

- Aboujaoude, E., Gega, L., & Saltarelli, A. J. (2021). The retention challenge in remote therapy and learning seen through the lens of the COVID-19 pandemic. *World Psychiatry*, 20(1), 138–139. <https://doi.org/10.1002/wps.20828>
- Ajzen, I. (2011). The theory of planned behaviour: Reactions and reflections. *Psychology & Health*, 26(9), 1113–1127. <https://doi.org/10.1080/08870446.2011.613995>
- Angelidis, A., Solis, E., Lautenbach, F., van der Does, W., & Putman, P. (2019). I’m going to fail! Acute cognitive performance anxiety increases threat-interference and impairs WM performance. *PLOS ONE*, 14(2), Article e0210824. <https://doi.org/10.1371/journal.pone.0210824>
- Bailenson, J. N. (2021). Nonverbal overload: A theoretical argument for the causes of zoom fatigue. *Technology, Mind, and Behavior*, 2(1). <https://doi.org/10.1037/tmb0000030>
- Blascovich, J. (2002). A theoretical model of social influence for increasing the utility of collaborative virtual environments. In *Proceedings of the 4th international conference on Collaborative virtual environments (CVE ’02)* (pp. 25–30). Association for Computing Machinery. <https://doi.org/10.1145/571878.571883>
- Blascovich, J., Mendes, W. B., Hunter, S. B., & Salomon, K. (1999). Social “facilitation” as challenge and threat. *Journal of Personality and Social Psychology*, 77(1), 68–77. <https://doi.org/10.1037/0022-3514.77.1.68>
- Bond, C. F. (1982). Social facilitation: A self-presentational view. *Journal of Personality and Social Psychology*, 42(6), 1042–1050. <https://doi.org/10.1037/0022-3514.42.6.1042>
- Bond, C. F., Jr., & 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>
- Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology*, 27(2), 270–295. <https://doi.org/10.1006/ceps.2001.1094>
- Claypoole, V. L., & Szalma, J. L. (2018). Facilitating sustained attention: Is mere presence sufficient? *The American Journal of Psychology*, 131(4), 417–428. <https://doi.org/10.5406/amerjpsyc.131.4.0417>
- Cottrell, N. B., Wack, D. L., Sekerak, G. J., & Rittle, R. H. (1968). Social facilitation of dominant responses by the presence of an audience and the mere presence of others. *Journal of Personality and Social Psychology*, 9(3), 245–250. <https://doi.org/10.1037/h0025902>
- Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, 130(3), 355–391. <https://doi.org/10.1037/0033-2909.130.3.355>
- 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>
- Elliot, A. J., Payen, V., Brisswalter, J., Cury, F., & Thayer, J. F. (2011). A subtle threat cue, heart rate variability, and cognitive performance. *Psychophysiology*, 48(10), 1340–1345. <https://doi.org/10.1111/j.1469-8986.2011.01216.x>
- Engert, V., Efanov, S. I., Duchesne, A., Vogel, S., Corbo, V., & Pruessner, J. C. (2013). Differentiating anticipatory from reactive cortisol responses to psychosocial stress. *Psychoneuroendocrinology*, 38(8), 1328–1337. <https://doi.org/10.1016/j.psyneuen.2012.11.018>
- Fiske, A., Henningsen, P., & Buys, A. (2019). Your robot therapist will see you now: Ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *Journal of Medical Internet Research*, 21(5), Article e13216. <https://doi.org/10.2196/13216>
- Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current Biology*, 17(16), R724–R732. <https://doi.org/10.1016/j.cub.2007.05.068>
- Garau, M., Slater, M., Pertaub, D.-P., & Razaque, S. (2005). The responses of people to virtual humans in an immersive virtual environment. *Presence: Teleoperators and Virtual Environments*, 14(1), 104–116. <https://doi.org/10.1162/1054746053890242>
- Gratch, J., Lucas, G., King, A., & Morency, L. P. (2014). It’s only a computer: The impact of human-agent interaction in clinical interviews. *Proceedings of the 14th international conference on autonomous agents and multiagent systems, AAMAS 2014* (Vol. 1, pp. 85–92). International Foundation for Autonomous Agents and Multiagent Systems. <https://dl.acm.org/doi/abs/10.5555/2615731.2615748>
- Green-Hamann, S., Campbell Eichhorn, K., & Sherblom, J. C. (2011). An exploration of why people participate in second life social support groups. *Journal of Computer-Mediated Communication*, 16(4), 465–491. <https://doi.org/10.1111/j.1083-6101.2011.01543.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](https://doi.org/10.1016/0022-1031(86)90040-5)
- Guerin, B., & Innes, J. M. (1984). Explanations of social facilitation: A review. *Current Psychological Research & Reviews*, 3(2), 32–52. <https://doi.org/10.1007/BF02686548>

- Hamilton, A. F. 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>
- Hoyt, C. L., Blascovich, J., & Swinb, K. R. (2003). Social inhibition in immersive virtual environments. *Presence: Teleoperators and Virtual Environments*, 12(2), 183–195. <https://doi.org/10.1162/105474603321640932>
- Jack, A. I., Dawson, A. J., Begany, K. L., Leckie, R. L., Barry, K. P., Ciccio, A. H., & Snyder, A. Z. (2013). fMRI reveals reciprocal inhibition between social and physical cognitive domains. *NeuroImage*, 66, 385–401. <https://doi.org/10.1016/j.neuroimage.2012.10.061>
- Jain, A., Bansal, R., Kumar, A., & Singh, K. D. (2015). A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students. *International Journal of Applied & Basic Medical Research*, 5(2), 124–127. <https://doi.org/10.4103/2229-516X.157168>
- Jin, S. A. A. (2010). The effects of incorporating a virtual agent in a computer-aided test designed for stress management education: The mediating role of enjoyment. *Computers in Human Behavior*, 26(3), 443–451. <https://doi.org/10.1016/j.chb.2009.12.003>
- Joinson, A. N. (2001). Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. *European Journal of Social Psychology*, 31(2), 177–192. <https://doi.org/10.1002/ejsp.36>
- Kothgassner, O. D., Goreis, A., Kafka, J. X., Kaufmann, M., Atteneder, K., Beutl, L., Hennig-Fast, K., Hlavacs, H., & Felnhofner, A. (2019). Virtual social support buffers stress response: An experimental comparison of real-life and virtual support prior to a social stressor. *Journal of Behavior Therapy and Experimental Psychiatry*, 63, 57–65. <https://doi.org/10.1016/j.jbtep.2018.11.003>
- Loveys, K., Fricchione, G., Kolappa, K., Sagar, M., & Broadbent, E. (2019). Reducing patient loneliness with artificial agents: Design insights from evolutionary neuropsychiatry. *Journal of Medical Internet Research*, 21(7), Article e13664. <https://doi.org/10.2196/13664>
- Monfardini, E., Redouté, J., Hadj-Bouziane, F., Hynaux, C., Fradin, J., Huguet, P., Costes, N., & Meunier, M. (2016). Others' sheer presence boosts brain activity in the attention (but not the motivation) network. *Cerebral Cortex*, 26(6), 2427–2439. <https://doi.org/10.1093/cercor/bhv067>
- Okita, S. Y., Bailenson, J., & Schwartz, D. L. (2007). The mere belief of social interaction improves learning. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 29, 1355–1360. <https://escholarship.org/uc/item/7rs81781>
- Park, S., & Catrambone, R. (2007). Social facilitation effects of virtual humans. *Human Factors*, 49(6), 1054–1060. <https://doi.org/10.1518/001872007X249910>
- Pfeiffer, U. J., Schilbach, L., Timmermans, B., Kuzmanovic, B., Georgescu, A. L., Bente, G., & Vogeley, K. (2014). Why we interact: On the functional role of the striatum in the subjective experience of social interaction. *NeuroImage*, 101, 124–137. <https://doi.org/10.1016/j.neuroimage.2014.06.061>
- 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>
- Platanias, J., & Moran, G. P. (2001). Social facilitation as a function of the mere presence of others. *The Journal of Social Psychology*, 141(2), 190–197. <https://doi.org/10.1080/00224540109600546>
- Rajecki, D. W., Ickes, W., Corcoran, C., & Lerner, 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>
- Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Center for the Study of Language and Information. Cambridge University Press.
- Riedl, R., Mohr, P., Kenning, P., Davis, F., & Heekeren, H. (2014). Trusting humans and avatars: A brain imaging study based on evolution theory. *Journal of Management Information Systems*, 30(4), 83–114. <https://doi.org/10.2753/MIS0742-1222300404>
- Ring, L., Shi, L., Totzke, K., & Bickmore, T. (2015). Social support agents for older adults: Longitudinal affective computing in the home. *Journal on Multimodal User Interfaces*, 9(1), 79–88. <https://doi.org/10.1007/s12193-014-0157-0>
- Schmitt, B. H., Gilovich, T., Goore, N., & Joseph, L. (1986). Mere presence and social facilitation: One more time. *Journal of Experimental Social Psychology*, 22(3), 242–248. [https://doi.org/10.1016/0022-1031\(86\)90027-2](https://doi.org/10.1016/0022-1031(86)90027-2)
- Steinmetz, J., Xu, Q., Fishbach, A., & Zhang, Y. (2016). Being observed magnifies action. *Journal of Personality and Social Psychology*, 111(6), 852–865. <https://doi.org/10.1037/pspi0000065>
- Sutskova, O., Senju, A., & Smith, T. J. (2020). *Impact of video-mediated online social presence and observance on cognitive performance* [Registered Report]. <https://doi.org/10.17605/OSF.IO/MBRZT>
- Tennie, C., Frith, U., & Frith, C. D. (2010). Reputation management in the age of the world-wide web. *Trends in Cognitive Sciences*, 14(11), 482–488. <https://doi.org/10.1016/j.tics.2010.07.003>
- Wellner, M., Sigrist, R., von Zitzewitz, J., Wolf, P., & Riener, R. (2010). Does a virtual audience influence rowing? *Proceedings of the Institution of Mechanical Engineers. Part P, Journal of Sports Engineering and Technology*, 224(1), 117–128. <https://doi.org/10.1243/17543371JSET33>
- 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/10.1126/science.149.3681.269>
- Zanbaka, C. A., Ullinski, 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* (pp. 1561–1570). <https://doi.org/10.1145/1240624.1240861>

(Appendices follow)

Appendix A

Analysis Plan

A copy of the preregistered analysis plan, followed in the Results section of the current report. For the full Registered Report see our OSF page (Sutskova et al., 2020).

Two separate ANOVAs with the three independent factors above were carried out to investigate RT and Accuracy. Simple effect analyses followed up the ANOVA to assess a priori planned predictions of hypotheses as specified below. These ANOVAs were then followed up with a series of planned contrast, to examine the direction of effects within each level of difficulty based on the expectation from SFE. The specific contrasts for each hypothesis are described below.

Hypothesis 1 would be supported by the presence of significant difficulty \times observance interaction. Hypothesis 1 predicted a main effect of observance for each level of difficulty, expected from SFE. According to SFE, in the easy task, the no observer condition should show “worse” performance (i.e., lower accuracy and slower RT) than the observer present condition (a combination of nonattentive and attentive conditions). In the difficult task, the no observer condition should show “better” performance (i.e., higher accuracy and faster RT) than nonattentive and attentive conditions. A priori contrasts were conducted between no observer conditions versus the combination of nonattentive and attentive observer conditions,

which were conducted separately at each difficulty level (easy, difficult).

Hypothesis 2 would be supported by the presence of significant difficulty \times observance \times conspecific interaction. The follow-up analysis for the three-way interaction consisted of planned contrasts within difficulty \times observance interaction, for each conspecific group separately. The planned contrast compared performance changes within each of the difficulty levels (easy, difficult), comparing attentive observer versus the not monitored condition (a combination of no observer and nonattentive observer conditions). In the easy task, the attentive observer condition should show “better” performance (i.e., higher accuracy and faster RT) than nonattentive and no observer conditions. In the difficult task, the attentive observer condition should show “worse” performance (i.e., lower accuracy and slower RT) than nonattentive and no observer conditions. We expected these results to be significant in real-human and avatar groups, but not in the agent group. The Bayesian analysis will be used to test support for the null effect in agency groups.

Exploratory Bonferroni corrected posthoc comparisons will be conducted to investigate further effects, such as the differences in the magnitude of SFE’s between all three conspecific conditions.

Appendix B

Deviations From Registered Report

There were a few minor deviations from the original Registered Report (Sutskova et al., 2020). Below, we list the sections which were affected and the reasoning behind the deviation from the original preregistered plan.

Participants

In the Registered Report, we originally predicted the recruitment of 70 participants, 54 required by the power analysis estimate, plus an additional 25% to account for the expected drop-out due to either exclusion criteria or technical errors. We intended to stop testing once the expected analysis sample size ($N = 54$) was acquired. Given the challenges of remote online testing and issues with using a fairly complicated video communication setup at participants homes, we ended up having to recruit 90 participants to reach the target analysis sample size ($N = 54$).

Design

The original report suggested the exclusion of participants who performed outside of 2 *SD* (95 percentile per difficulty) of the

average data range. However, due to an unexpected high participant drop off rate relating to remote testing during the COVID-19 pandemic (and considering the already additional testing of 20 participants), we decided to include participants within the 3 *SD* of the dataset range, making sure our analysis has predicted power. Analysis of variance (see Results) showed a fairly even data distribution between groups, suggesting that 3 *SD* inclusion did not affect our sample significantly.

Analysis Plan

The preregistered plan was to run an additional analysis using a diffusion drift model, as an explorative analysis alongside the main ANOVA analyses. Since this analysis was merely exploratory and did not contribute directly to any of the hypotheses postulated, the analysis was dropped from this manuscript for the sake of space and cohesion.

We have changed the name for the merged level of “not attended” (nonattentive + no observer combined) used in the AE analysis to “not monitored” to avoid any terminological confusion between the “nonattentive” level of observance IV.

Appendix C

Additional Results

Additional nonsignificant results of the three-way analysis, which were not part of the postulated hypothesis. Nonsignificant results from conducted three-way ANOVA: 3 conspecific \times 3 observance \times 2 difficulty.

Percent Accuracy

Observance \times conspecific type: $F(4, 51) = 1.78$, $p = .14$, $\eta_p^2 = .065$.

Difficulty \times conspecific type: $F(2, 51) = 2.34$, $p = .11$, $\eta_p^2 = .084$.

Reaction Times

Observance \times conspecific type: $F(4, 51) = 0.25$, $p = .91$, $\eta_p^2 = .01$.

Difficulty \times conspecific type: $F(2, 51) = .001$, $p = .99$, $\eta_p^2 < .0001$.

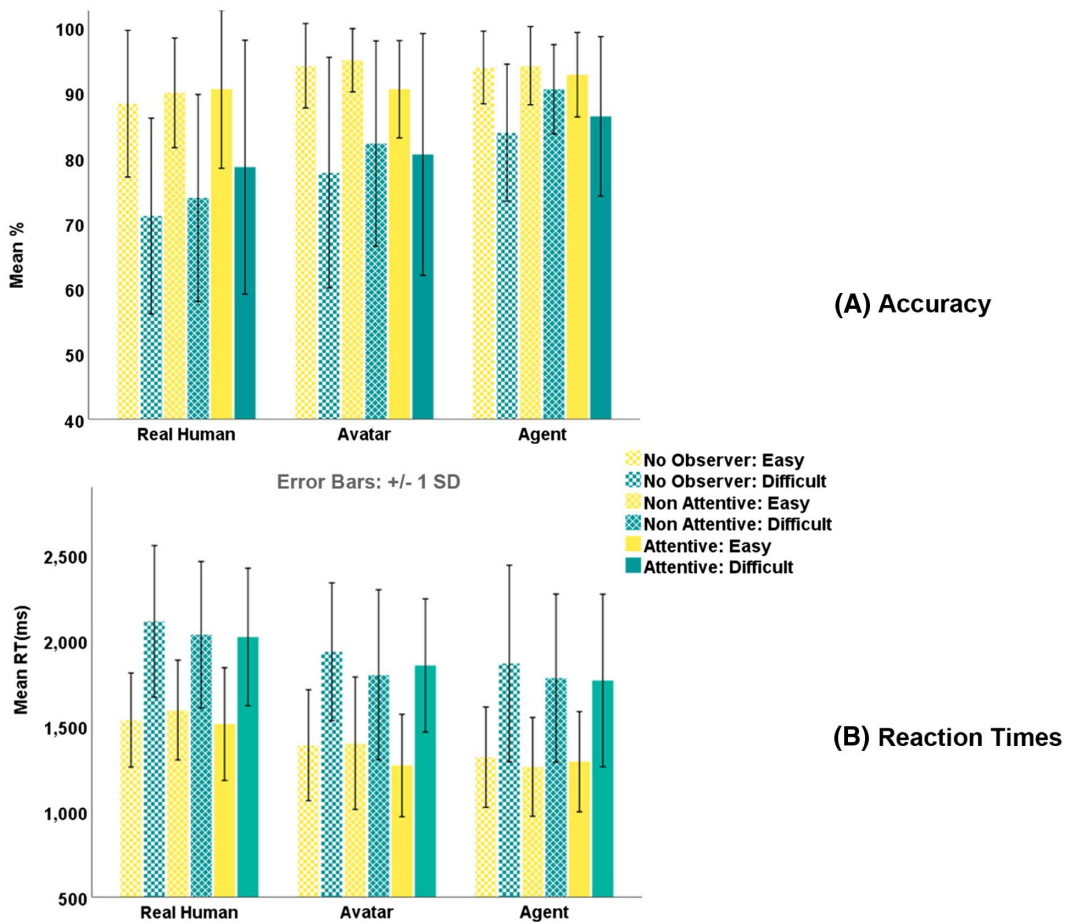
Appendix D

Breakdown of a Three-Way ANOVA

Additional figures represent means and 1 SD per each three-factor ANOVA level separately.

Figure D1

(A) and (B) Illustrate a Breakdown for All Conditions Means (M) and 1 Standard Deviation (SD) of the Three-Way ANOVA: Conspecific \times 3 Observance \times 2 Difficulty, for Accuracy (A) and RT (B) Separately



Note. ANOVA = analysis of variances.

(Appendices continue)

Appendix E

Breakdown of the Audience Effects (AE) Statistics for Each Conspecific Condition Separately

A series of planned *t*-test comparisons for the audience effects (AE) Hypothesis 2, with a mean (*M*) and 1 standard deviation (*SD*).

Table E1

Audience Effect, Percent Accuracy Descriptives (M and SD), and t and p Statistics for Planned (Not Corrected) Follow Up Contrasts, Within Each Conspecific Group Separately

Conspecific	Difficulty	Not monitored <i>M</i> (<i>SD</i>)	Attentive observer <i>M</i> (<i>SD</i>)	<i>t</i> (1, 17)	<i>p</i>
Real human	Easy	89.17 (8.91)	90.56 (12.11)	0.67	.51
	Difficult	72.50 (13.88)	78.61 (19.46)	2.02	.06
Avatar	Easy	94.58 (4.04)	90.56 (7.45)	1.91	.07
	Difficult	80.00 (15.58)	80.56 (18.54)	0.23	.82
Agent	Easy	94.02 (3.85)	92.78 (6.46)	0.71	.49
	Difficult	87.22 (6.91)	86.39 (12.22)	0.35	.73

Note. Marginal effects for real human difficult ($p = .06$) and avatar easy conditions ($p = .07$), with statistical significance cutoff estimation of $p < .05$.

Table E2

Audience Effect, Reaction Times Descriptives (M and SD), and t and p Statistics for Planned (Not Corrected) Follow Up Contrasts, Within Each Conspecific Group Separately

Conspecific	Difficulty	Not monitored <i>M</i> (<i>SD</i>)	Attentive observer <i>M</i> (<i>SD</i>)	<i>t</i> (1, 17)	<i>p</i>
Real human	Easy	1564.39 (267.80)	1511.06 (329.45)	0.89	.39
	Difficult	2073.27 (402.02)	2020.95 (402.55)	0.79	.44
Avatar	Easy	1393.57 (320.62)	1269.01 (299.76)	2.68	.016*
	Difficult	1867.36 (397.15)	1854.19 (390.13)	0.20	.84
Agent	Easy	1289.51 (265.35)	1291.23 (293.29)	0.04	.97
	Difficult	1823.14 (513.30)	1766.64 (504.78)	0.76	.46

Note. Significant effects for avatar easy conditions (* $p = .016$), with statistical significance cutoff estimation of $p < .05$. Bold values represent the effects which is non-significant.

Received August 10, 2020
Revision received November 29, 2021
Accepted December 5, 2021 ■