Fake People, Real Effects

The Presence of Virtual Onlookers Can Impair Performance and Learning

Wouter Durnez[1(✉)], Klaas Bombeke[1], Jamil Joundi[1], Aleksandra Zheleva[1], Emiel Cracco[2], Fran Copman[2], Marcel Brass[2], Jelle Saldien[1], and Lieven De Marez[1]

[1] imec-mict-UGent, Department of Communication Sciences, Ghent University, Miriam Makebaplein 1, 9000 Gent, Belgium
wouter.durnez@ugent.be

[2] Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Gent, Belgium

Abstract. Can effects of social influence be elicited in virtual contexts, and if so, under which conditions can they be observed? Answering these questions has theoretical merit, as the answers can help broaden our understanding of the interaction mechanisms described by social psychology. The increasing popularity of immersive media in training applications, however, has made these questions of practical significance. Virtual reality (VR), in particular, is a weapon of choice in designing training and education simulations, as it can be used to generate highly realistic characters and environments. As a consequence, it is key to understand under which circumstances virtual 'others' can facilitate or impede performance and – especially – learning. In this study, we investigated the impact of virtual onlookers on an adapted Serial Reaction Time (SRT) task that was presented in VR. In each trial, participants responded to a series of spherical stimuli by tapping them with handheld controllers when they lit up. Depending on the experiment block, the sequence order was either the permutation of a *fixed* order (and therefore predictable given the first stimulus), or fully *random* (and therefore unpredictable). Participants were divided into three groups (*audience* variable), depending on the environment in which the task was set: a group without onlookers (*none* condition), a group with a computer-generated audience (*CGI* condition), and a group being watched by a prerecorded audience (*filmed* condition). Results showed that the presence of a virtual audience can hamper both overall performance and learning, particularly when the audience appears more realistic. This study further reinforces the notion that the effects of social influence transcend the physical presence of others, but rather extend to virtual audiences.

Keywords: Social inhibition · Social facilitation · Learning · Virtual reality

W. Durnez, K. Bombeke and J. Joundi–These authors contributed equally to this work.

© Springer Nature Switzerland AG 2020
J. Y. C. Chen and G. Fragomeni (Eds.): HCII 2020, LNCS 12190, pp. 440–452, 2020.
https://doi.org/10.1007/978-3-030-49695-1_29

1 Introduction

In the presence of others, we behave differently [12]. Consider the following scenarios. In the first, a musician puts on the recital of a lifetime, encouraged by the opportunity to showcase his skills in front of admiring spectators. In the second, the mind of a young researcher goes blank as he – in spite of his numerous flawless repetitions the day before – stumbles his way through a conference presentation. These so-called audience effects are examples of what Zajonc [15] called "the oldest experimental paradigm of social psychology" - *social facilitation*.

The *social facilitation effect* (SFE) refers to the finding that, when observed by other people, an individual's performance will be better on simple and well-trained tasks, yet worse on complex and new tasks [2,15]. Several theories have attempted to explain this phenomenon, of which Zajonc's social facilitation theory (or generalized drive hypothesis) sparked the most interest [15,16]. The theory argues that the mere presence of others will increase arousal, facilitating dominant responses or automatic reflexes. In simple, well-learned tasks, these dominant responses will enhance performance (i.e., social facilitation), whereas in complex, novel tasks, they need to be overruled by more appropriate responses, impairing performance (also called "social inhibition") [15,16].

Social facilitation has since been the focus of many research endeavors [12,13]. With the rapid progression of computers and computer-generated images (CGI), however, an interesting new research angle presented itself. What if the 'other' was no longer physically present, but represented through a virtual avatar? What if the virtual avatar was controlled, not by another human being, but by a simple computer program? A study by Rickenberg and Reeves [11] suggested that the psychology of human-human relationships translates to how people react to virtual agents, as the visual properties of these avatars automatically trigger social responses. With respect to the SFE, however, the evidence has been largely inconsistent [1,8,10,17].

Hoyt and colleagues [8] conducted a study in which participants performed two categorization tasks of different difficulty levels. They either did so alone, or in front of a virtual audience. Participants in the latter condition were either led to believe that the audience were human-controlled avatars, or computer-controlled agents. Interestingly, the experimenters only found evidence for *social inhibition* in the *avatar* condition, suggesting that the driving force in the effect was not the puppet, but who (or what) the participants believed to pull its strings. No other social facilitation (or inhibition) effects were found. A similar result was obtained by Zanbaka and colleagues [17]. They found evidence for social inhibition only when female participants performed novel categorization and pattern recognition tasks in the presence of a virtual human. No such effects were found among the male participants.

More recent studies have failed to paint a clearer picture. In 2007, Park and others [10] found evidence for both social facilitation and social inhibition for a variety of tasks that were performed in the presence of a virtual agent. Notably, these effects were found in spite of informing participants that the virtual deuteragonist was an artificial intelligence, and therefore not human-controlled. A 2010 study by Hayes and colleagues [6] found trends towards social facilitation and inhibition in performance of simple and complex math tasks. However, none of the effects reached significant levels. Furthermore, they showed that the gender of the observer impacted the participants' sense of presence ratings. Nonetheless, no straightforward explanation was offered regarding this effect. Baldwin and colleagues, finally, conducted three different experiments that were specifically geared towards replication of the SFE, and failed to find any evidence [1].

Why does it seem so hard to converge on a set of replicable results? Although there can be many possible explanations, these mixed findings can arguably be attributed – in part – to problems with the design and difficulty level of different experimental tasks. More specifically, a common problem in social facilitation research relates to the ceiling effect present in the simple task (see also [17]) but absent in the complex task. Typically, participants first undergo a practice phase, after which they have to meet a certain accuracy criterion before they move on to the main blocks. As a result, participants are already highly skilled at the onset of the actual experimental phase, leaving little to no room for improvement. In other words, the simplicity of the task makes it impossible to detect additional effects of the presence of an audience. A second problem relates to the difference in task difficulty between the simple and the complex task itself. By only having the rather arbitrarily chosen dichotomy between simple and complex tasks, it is possible that previous studies failed to observe more subtle social facilitation and inhibition effects.

Furthermore, the aforementioned research efforts neglected to take the level of realism of their avatars into account, using virtual entities ranging from Microsoft Word's Clip (e.g., [5]) to more realistic models of human beings (e.g., [1,17]). To the best of our knowledge, there has been no previous research comparing the effects that virtual avatars of different levels of realism have on behavior. Given today's technology, however, virtual humans in three-dimensional environments can be made very similar to real humans. 360° video, for instance, makes it possible to use actual humans as observers in the virtual environment by using prerecorded footage. Comparing the effects of a virtual audience consisting of real (albeit filmed) people with an audience of artificially crafted 3D avatars would help elucidate the role of the audience's visual credibility. Is it enough that we are being watched by other 'agents', or is it necessary that we – on some level – perceive the audience as 'realistic human beings'?

In this study, we aimed to address the shortcomings of earlier research by introducing a VR-tailored serial reaction time (SRT) task. In this task, participants had to respond to a sequence of illuminated spheres as quickly and accurately as possible. Depending on the condition participants were assigned to, they performed this task in front of a CGI audience, a filmed audience, or in an empty room. In addition, in half of the experimental blocks the sequence was made predictable given the first lit up sphere. Participants could learn an underlying sequence, which enabled them to anticipate the illumination of the next sphere, allowing them to (potentially) respond more quickly and more accurately.

This task has two main advantages compared to the studies discussed earlier. First, it is less likely that we come across a ceiling effect, given the main dependent measure: reaction time. Second, by introducing a memory element, we are able to – for the first time – investigate the effects of (virtual) onlookers on learning speed. If participants succeed in memorizing the underlying sequence in fixed order blocks, this should be reflected in lower response times and higher accuracy scores. In doing so, we avoid the typical (and somewhat arbitrary) dichotomy between a simple and a complex task, focusing on learning performance instead. Besides the aforementioned methodological advantages, the focus on learning processes instead of final performance makes findings with regard to social facilitation and inhibition effects more ecologically valid and more informative for real-world applications, given the increasing ubiquitousness of immersive media in the fields of training and education. For example, in the manufacturing industry, virtual and augmented reality are increasingly used to train new employees, with employers trying to find the most efficient and effective ways to do so. Our study thus aims to inform developers of such training environments on the impact of the virtual 'people' they populate their simulators with.

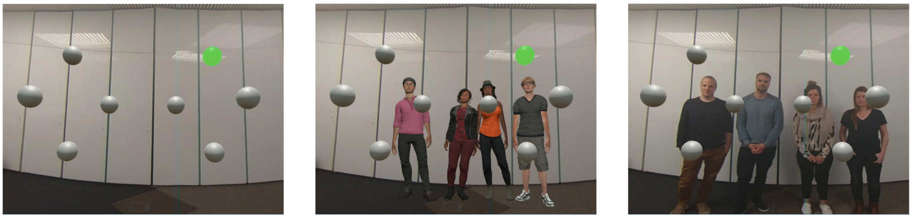
2 Method

2.1 Participants

For this study, 60 participants with normal or corrected eyesight were recruited through online sampling. Due to technical difficulties at the onset of the experiment, the first participant's data were omitted from the dataset. Our participant pool consisted of more women (76.3%; M = 26.2 years old; SD = 8.9 years old) than men (23.7%; M = 31.4 years old; SD = 13.9 years old). Almost all of them were highly educated or still attending university (52.5% had a bachelor's degree, and 22% had a master's degree). All subjects participated on a voluntary basis and signed an informed consent with ethical approval from the university's ethical committee.

2.2 Design and Stimulus Material

Using Unity (version 3.5), we constructed a custom-made environment in which our SRT task was performed. Participants were assigned to one of three groups (*audience* variable): a group performing the task in an empty (virtual) room (*none* condition), a group performing the task in the presence of a CGI audience (*CGI* condition), and a group performing the task in the presence of a filmed audience (*filmed* condition). They were sat down in our lab and the experimenter put on the head-mounted display (HMD). Doing so, each participant found themselves in the same (virtual) room: a co-creation space that was recorded using a 360° video camera. In the *none* condition, this room was empty (Fig. 1a). Participants in the *CGI* condition were placed across four 3D-rendered virtual characters (Fig. 1b). These characters were superimposed on the room in such a manner that their presence seemed natural (i.e., they were scaled appropriately and appeared to stand on the floor, rather than float). People in the *filmed* condition, finally, saw four 'real' people in front of them – their 360° backdrop video was one of the same room, yet now containing four actual people looking at the camera (Fig. 1c). Importantly, we made sure that each audience contained an equal number of male and female onlookers, so effects could not be attributed to gender differences between them [6]. Each video only lasted about 10 s, but was looped in such a way that they contained no video glitches (e.g., abrupt changes in movement or facial expressions caused by cutting the original video). To achieve this fluency, the videos were copied, reversed, and concatenated, resulting in a video that could be repeated indefinitely.



(a) *None* (b) *CGI* (c) *Filmed*

Fig. 1. Snapshot of the virtual environments corresponding to each of the *audience* conditions. (Color figure online)

In our experiment, we used an adapted version of the serial reaction time task (SRT). In front of each participant, 8 grey spheres floated in the air. They were positioned in two diamond formations, 4 on the left and 4 on the right of the participant. We made sure that the spheres were spread out in order to minimize the chance of accidentally hitting the wrong sphere. The formation of the spheres, in turn, was chosen in an effort to contain all spheres within the participant's (HMD-imposed) field of view. In other words, if a participant was looking straight ahead, all spheres were visible. To ensure this was the case for each participant, the experimenters made sure that everyone sat in exactly the same spot and did not move their chair. A single trial in the task was as follows: 8 spheres lit up (turned fluorescent green, see Fig. 1) in a certain sequence. Participants were asked to tap the illuminated sphere using one of two controllers, one of which was placed in each hand. Tapping a sphere caused it to extinguish (turn grey again), and the next sphere to illuminate. Participants felt a vibration in the respective controller when tapping a sphere, irrespective of whether they tapped the correct (green) sphere. Each illumination lasted for 5 s or until a response was registered, giving participants ample time to respond. However, participants were asked to respond as fast and as accurate as possible. Participants were given no instructions on which controller to use for which of the spheres, and were free to choose their approach. When piloting, we noticed that some responses were incorrectly registered as 2 separate taps, due to the irregular geometry of the controller meshes colliding with the spheres. These ‘double taps’ had unrealistically fast response times (i.e., under 100 ms). To mitigate this issue, we set an inactivation period of 150 ms after each tap. Spheres could already illuminate in this window, but responses were not registered. Piloting confirmed that the response time of intentional reactions always exceeded this window, leaving participants oblivious to this technical restriction.

All data was logged in the Unity application. For each trial, we recorded *audience* (none, CGI, or filmed), *order* (random or fixed), trial number, block number (both overall and within each level of the *order* variable, e.g., the 3th random block), the exact order in which the 8 spheres were illuminated during that trial, and the reaction time (RT) to respond to each single sphere illumination. Accuracy was also logged on a sphere-level – if a participant erroneously tapped a sphere that was not illuminated, this was counted as an incorrect response. A correct trial consisted of 8 correct responses (tapped spheres). Any mistake caused the trial to be labeled as incorrect.

2.3 Procedure and Technical Set-Up

Upon arrival in the lab, participants were briefed on the experiment and what was expected of them. They signed an informed consent and were asked to take place on a chair in the middle of the room, next to a large television screen on which the experimenter could see what the participant observed in VR. Next, the experimenter carefully put the HTC Vive HMD on the participants’ head and two HTC Vive controllers in their hands. The HTC Vive HMD offers a resolution of 2160 $\times$ 1200 (with 1080 $\times$ 1200 per eye), global lighting and AMOLED-displays of 90 Hz. After making sure the participants’ vision was not blurry because of incorrect placement, the experiment started with an instruction screen.

Participants were first presented with a practice block of 14 trials. In this block, the sequences were completely random – the illumination pattern could thus not be predicted. Next, participants completed 8 blocks of 14 trials, resulting in a total of 112 trials. In half of these blocks, the sequence was random. In the other half, the sequences were permutations of a fixed base sequence (e.g., given the base sequence of 1-2-3-4-5-6-7-8, a first sequence could be 3-4-5-6-7-8-1-2, followed by 5-6-7-8-1-2-3-4, etc.). In other words, participants could *learn* the underlying sequence, after which all sphere illuminations were predictable given the first illuminated sphere. Importantly, participants were informed of this mechanism, allowing them to try and learn the pattern from the start. Block order alternated, and was balanced (i.e., the first block was fixed or random an equal number of times over all participants). Participants were also informed about the block type (*order*: random vs. fixed) at the start of each block.

2.4 Analysis

We analyzed response time (RT) and (trial) accuracy using linear mixed-effects models with a Gaussian link function and a binomial link function, respectively. All statistical modeling consisted of the following steps. First, we entered the between-subject variable *audience*, the within-subject variables *order* and *block number* as well as all interactions (to the third degree) in the model as fixed factors. We also added a random effect to reflect adjustments to the intercept, conditional on the subject variable. We then verified whether the addition of a random effect for the *order* and *block number* variables, conditional on *subject*, increased the model's goodness of fit. If this was the case, the random effect was retained in the next modeling steps. In a second step, we sought out the most parsimonious model to fit the data, by systematically omitting non-significant fixed effects from the model. Models were compared using likelihood-ratio tests. The third and final step consisted of inspecting the best model's analysis of variance table and evaluating specific hypotheses. Significant interactions were investigated using post-hoc contrast analyses, which were corrected for multiple testing according to the Holm-Bonferroni method [7]. See [3,4,14] for a similar approach.

3 Results

3.1 Reaction Time

To analyze the RT data, we first omitted all incorrect trials from the data set (i.e., trials where one or more of the spheres were not tapped in time, or where a wrong sphere was tapped instead). We then filtered the data by applying the interquartile range criterion to the RT data distribution. More explicitly, we calculated the interquartile range ($IQR = Q_3 - Q_1$) and removed RTs below $Q_1 - (1.5 \times IQR)$, as well as RTs over $Q_3 + (1.5 \times IQR)$. This resulted in a removal of 0.036% of all trials. Overall RT performance is visualized in Fig. 2, whereas the evolution of RT over time is shown in Fig. 3.

The model that best described the RT data contained all fixed effects and interactions, as well as a subject-based random intercept, a random effect of *block number*, and a random effect of *order*. In it, 2 main effects were significant (*order*: $\chi^2(1) = 32.886; p < 0.001$, and *block number*: $\chi^2(3) = 66.769; p < 0.001$), as well as 2 two-way interaction effects (*audience* x *block number*: $\chi^2(6) = 16; p < 0.001$, and *order* x *block number*: $\chi^2(3) = 208,842; p < 0.001$).

Since it contained a significant three-way interaction effect ($\chi^2(6) = 41.229; p < 0.001$), the model was not restricted further. However, in order to disentangle this complex interaction, the data was split up according to the levels of the *audience* variable. Statistical modeling was then reapplied to each subset.

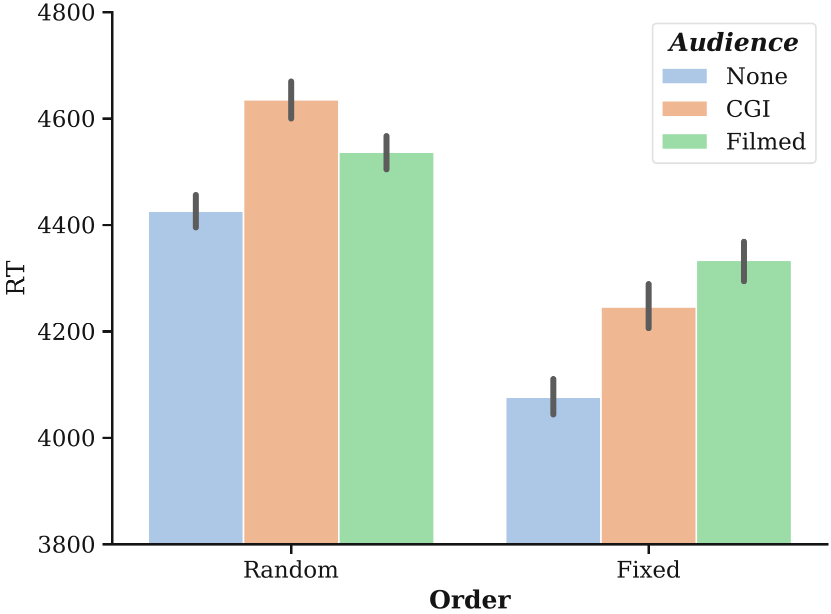


Fig. 2. Mean RT for blocks with random order sequences (left) and blocks with fixed order sequences (right). Bars reflect 95% confidence interval.

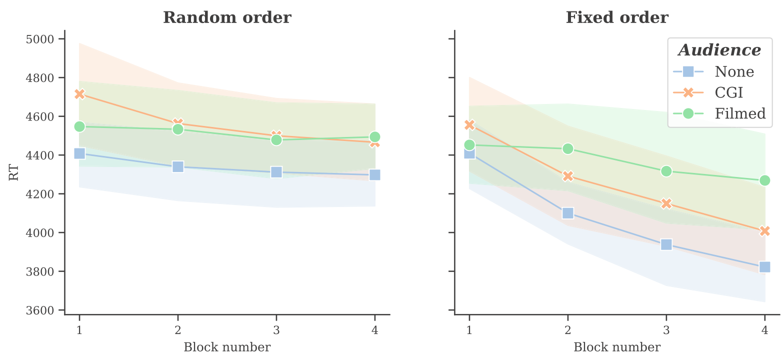


Fig. 3. Evolution of RT for blocks with random order sequences (left) and blocks with fixed order sequences (right). Bands reflect 95% confidence interval.

For the *none* group, the best model contained a subject-based random intercept, a random effect of *block number*, and a random effect of *order*. The model showed two significant main effects (*order*: $\chi^2(1) = 13.341; p < 0.001$; *block number*: $\chi^2(3) = 31.609; p < 0.001$), as well as a significant interaction ($\chi^2(3) = 41.229; p < 0.001$). Follow-up contrast analysis showed no significant improvement between the first and fourth random block. In the fixed blocks, however, RTs significantly decreased ($M_{block1} = 4410.509 \pm 534.726, M_{block4} = 3831.206 \pm 491.945, \chi^2(1) = 74.826, p < 0.001$).

In the *CGI* group, the best fit contained the same random and fixed effect structure. It too showed two significant main effects (*order*: $\chi^2(1) = 28.954; p < 0.001$; *block number*: $\chi^2(3) = 50.788; p < 0.001$) and a significant interaction ($\chi^2(3) = 56.797; p < 0.001$). Interestingly, follow-up contrasts suggested significant improvement between both the first and last fixed block ($M_{block1} = 4564.087 \pm 670.601, M_{block4} = 4018.998 \pm 614.880, \chi^2(1) = 66.961, p < 0.001$), and – to a lesser degree – between the first and last random block ($M_{block1} = 4713.209 \pm 673.210, M_{block4} = 4463.901 \pm 567.516, \chi^2(1) = 14.234, p < 0.001$).

The *filmed* group's data, finally, was best described by a model containing the same random effect structure as before, and a fixed term for *block number*. Both *order* and its interaction with *block number* were not significant, and were omitted from the model. The *block number* main effect was significant ($\chi^2(3) = 20.012, p < 0.001$). Following up on the previous contrast analyses, we also compared the RT in the first and last blocks (of fixed and random blocks combined), but found no significant difference ($M_{block1} = 4473.089 \pm 583.862, M_{block4} = 4347.862 \pm 583.519), \chi^2(1) = 3.400, p = 0.065$).

Since the previous results suggest that learning mostly occurred in the fixed blocks, we decided to compare the RT evolution in fixed blocks between different *audience* groups. To this effect, we restricted the data to contain only fixed blocks, and fit a separate model. This model contained a random subject-based intercept and a random effect of *block number*, as well as fixed factors for *audience* and *block number*, and their interaction. In the model, the interaction effect was significant ($\chi^2(6) = 15.111, p = 0.019$). We further explored this interaction through contrast analysis by comparing the difference between the first and fourth blocks (reflecting the RT gain – or learning rate) between groups. We found that the block-based RT evolution differed significantly between *CGI* and *filmed* groups ($\chi^2(1) = 7.737, p = 0.011$), and more so between the *none* and *filmed* groups ($\chi^2(1) = 9.139, p = 0.008$). The difference in learning rate between the *none* and *CGI* groups was not signficant.

3.2 Accuracy

The accuracy measure – reflecting whether all spheres were tapped in the right order (correct), or whether one or more mistakes were made (incorrect) – was modeled using the same approach (see Fig. 4). In this case, however, the best model contained a single fixed effect for the *audience* variable, a random subject-based intercept, and a random effect for the *order* variable. The effect of *audience* was significant ($\chi^2(2) = 6.343, p = 0.042$), showing higher accuracy for the *none* ($M = 0.926 \pm 0.262$), and *CGI* ($M = 0.937 \pm 0.243$) groups compared to the *filmed* group ($M = 0.882 \pm 0.262$).

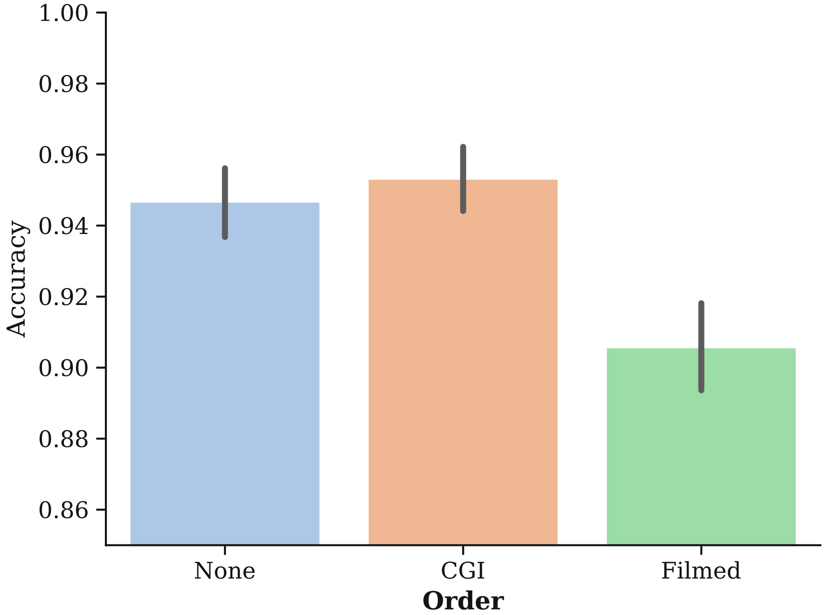


Fig. 4. Mean accuracy for *audience* conditions. Bars reflect 95% confidence interval.

4 Discussion

The aim of this study was to investigate the impact of virtual onlookers on performance and learning. To this effect, we designed a virtual environment in which participants performed an adapted SRT task by responding to a sequence of illuminating spheres. In half of the experimental blocks, the presented sequences were strictly random. In the other blocks, the sequences were – given the first illuminated sphere – predictable. In these blocks, participants could attempt to remember the underlying activation pattern, allowing them to anticipate the illumination of the next spheres and, thus, to perform better (i.e., faster reaction time, higher accuracy). As such, we were able to quantify each participant's learning rate, apart from overall performance.

Analyzing the participants' reaction time data clearly indicated that our adapted paradigm was effective, as participants were able to improve their response times on predictable (fixed) trials in both the *none* group (no audience) and the *CGI* group (artificially generated audience). Interestingly, the performance of participants in the *filmed* audience group did not improve significantly between the first and last fixed block, suggesting that their learning rate was hampered by the presence of the audience. In addition, the gain in response time was larger in the *none* group compared to the *CGI* group. This suggests that social inhibition may also have been at work in the *CGI* group, although we were unable to demonstrate this on a statistically significant level. Curiously, participants in the CGI group seemed to improve their response time to random trials over the course of the experiment, which could reflect increasing skill in the motor aspect of our task (e.g., handling the controllers with increasing precision). Further inspection of the data suggested that this was due to an especially poor performance in the first block (see Fig. 3), rather than an increased learning speed. As such, we are hesitant to attribute these observation to a potential effect of the *CGI* audience.

The results of our analyses on the accuracy data showed that, on average, people in the *filmed* group made more mistakes than those in other groups. Participants in the *CGI* and *none* groups were equally accurate. Again, these results indicate that virtual audiences can inhibit learning, not only in terms of response speed, but also in terms of accuracy. Furthermore, this seems to especially be true when these audiences are photo-realistic.

This study is, to the best of our knowledge, the first to present evidence of social inhibition caused by a virtual audience in a virtual environment. This evidence was by far most convincing when the audience consisted of recorded people. Still, we also found indications that CGI agents can influence performance. Together, these findings indicate that virtual agents can affect behavior – in this case: impair learning – but that the level of visual realism matters. In that respect, it is also interesting to note that participants were at all times aware that the environments and the characters therein were not human-controlled, but were either computer-generated or prerecorded. As such, our findings can hardly be attributed to their belief in human-controlled avatars, contrasting our results with those of Hoyt and colleagues [8].

This study has clear practical relevance in this era of immersive media, where VR-based training applications are increasingly prevalent. As computer graphics become more photo-realistic, social effects are more likely to arise. Specifically, our study predicts that photo-realistic audiences are likely to impair the performance and learning rate of those who attempt to execute tasks in virtual training environments. In that regard, two additional remarks present themselves.

First, it is likely that the correlation between the strength of a social effect and the visual realism of the virtual agents is not strictly linear. Earlier research by Mori and colleagues [9] uncovered the existence of the so-called 'uncanny valley effect': the observation that highly realistic yet imperfect human-like avatars are appraised more negatively compared to their ostensibly less realistic counterparts. In other words, unless the CGI representation is absolutely "perfect", people will notice subtle abnormalities in the representation which might result in an adverse response. Readers should thus be cautioned in extrapolating our findings to different degrees of 'realism'.

Second, it may be tempting to interpret the results of this study as a warning not to include realistic 'others' in virtual training environments. Indeed, doing so might impair the user's performance and learning speed, which at first glance may seem undesirable. However, it is important to note that our study did not evaluate performance beyond the contexts participants were assigned to. What would happen to task performance when people who previously practiced the task in an empty environment are made to execute it in front of an audience?

Could it be that they will be less capable to deal with the effect of the audience? Conversely, are people who trained in front of an audience better armed to deal with their influence? It is conceivable that, while training in front of (virtual) audiences slows down learning, it also prepares the trainees to handle the social impact of onlookers. A surgeon who only practices a procedure in an empty virtual environment may feel daunted having to perform it in front of fellow surgeons for the first time, partially negating the training effects of the simulation. Future experiments should be conducted to evaluate the cross-over effects of training in (populated) virtual environments. Regardless of the outcome of such studies, the present research makes one thing apparent: the onlookers in a virtual environment may be fake, but their social effects can be felt all the same.

References

1. Baldwin, N., Branyon, J., Sethumadhavan, A., Pak, R.: In search of virtual social facilitation effects. Proc. Hum. Factors Ergon. Soc. **59**(1), 90–94 (2015)
2. Bond, C.F., Titus, L.J.: Social facilitation: a meta-analysis of 241 studies. Psychol. Bull. **94**(2), 265 (1983)
3. De Paepe, A.L., Crombez, G., Legrain, V.: Remapping nociceptive stimuli into a peripersonal reference frame is spatially locked to the stimulated limb. Neuropsychologia **101**(May), 121–131 (2017). https://doi.org/10.1016/j.neuropsychologia.2017.05.015
4. Durnez, W., Van Damme, S.: Trying to fix a painful problem: the impact of pain control attempts on the attentional prioritization of a threatened body location. J. Pain **16**(2), 135–143 (2015)
5. Hall, B., Henningsen, D.D.: Social facilitation and human-computer interaction. Comput. Hum. Behav. **24**(6), 2965–2971 (2008). https://doi.org/10.1016/j.chb.2008.05.003
6. Hayes, A.L., Ulinski, A.C., Hodges, L.F.: That Avatar is looking at me! social inhibition in virtual worlds. In: Allbeck, J., Badler, N., Bickmore, T., Pelachaud, C., Safonova, A. (eds.) IVA 2010. LNCS (LNAI), vol. 6356, pp. 454–467. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15892-6_49
7. Holm, S.: A simple sequentially rejective multiple test procedure. Scand. J. Stat. **6**(2), 65–70 (1979)
8. Hoyt, C.L., Blascovich, J., Swinth, K.R.: Social inhibition in immersive virtual environments. Presence Teleoper. Virtual Environ. **12**(2), 183–195 (2003). https://doi.org/10.1162/105474603321640932
9. Mori, M., MacDorman, K.F., Kageki, N.: The uncanny valley. IEEE Rob. Autom. Mag. **19**(2), 98–100 (2012). https://doi.org/10.1109/MRA.2012.2192811
10. Park, S., Catrambone, R.: Social facilitation effects of virtual humans. Hum. Factors J. Hum. Factors Ergon. Soc. **49**(6), 1054–1060 (2007). https://doi.org/10.1518/001872007x249910
11. Rickenberg, R., Reeves, B.: The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. In: Proceedings of the SIGCHI conference on Human factors in computing systems - CHI 2000, vol. 2, no. 1, pp. 49–56 (2000). https://doi.org/10.1145/332040.332406
12. Sanders, G.S.: Driven by distraction : an integrative review of social facilitation theory and research. J. Exp. Soc. Psychol. **17**(3), 227–251 (1981)

13. Strauss, B.: Social facilitation in motor tasks: a review of research and theory. Psychol. Sport Exerc. **3**(3), 237–256 (2002). https://doi.org/10.1016/s1469-0292(01)00019-x
14. Verbruggen, F., Aron, A.R.: Theta burst stimulation dissociates attention and action updating in human inferior frontal cortex. Proc. Natl. Acad. Sci. **107**(31), 13966–13971 (2010). https://doi.org/10.1073/pnas.1001957107/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1001957107, http://www.pnas.org/content/107/31/13966.short
15. Zajonc, R.B.: Social facilitation. Science **149**(3681), 269–274 (1965). https://doi.org/10.1017/cbo9780511628214
16. Zajonc, R.B.: Feeling and thinking: preferences need no inferences. Am. Psychol. (1980). https://doi.org/10.1037/0003-066x.35.2.151
17. Zanbaka, C.A., Ulinski, A.C., Goolkasian, P., Hodges, L.F.: Effects of virtual human presence on task performance. In: Proceedings Artificial Reality and Telexistence, pp. 174–181 (2004)