Check for updates

The Impact of Self-Representation and Consistency in Collaborative Virtual Environments

Tara Collingwoode-Williams*, Zoë O'Shea, Marco Gillies and Xueni Pan

Department of Computing, Goldsmiths, University of London, London, United Kingdom

This paper explores the impact of self-representation (full body Self Avatar vs. Just Controllers) in a Collaborate Virtual Environment (CVE) and the consistency of self-representation between the users. We conducted two studies: Study 1 between a confederate and a participant, Study 2 between two participants. In both studies, participants were asked to play a collaborative game, and we investigated the effect on trust with a questionnaire, money invested in a trust game, and performance data. Study 1 suggested that having a Self Avatar made the participant give more positive marks to the confederate and that when the confederate was without an avatar, they received more trust (measured by money). Study 2 showed that consistency led to more trust and better productivity. Overall, results imply consistency improves trust only when in an equal social dynamic in CVE, and that the use of confederate could shift the social dynamics.

Keywords: virtual reality, human computer interaction, collaborative virtual environment, avatar representation, embodied consistency

1. INTRODUCTION

Collaborate Virtual Environments (CVE) can be used effectively in a multitude of different industries; more commonly applicable are those that utilise virtual reality (VR) for training, education, and entertainment. The advantage of a CVE is that it allows for interactions and controlled conditions that would not be possible in real life. For example, the ability to be virtually present in the same environment as someone that lives across the world is the fundamental feature that many social VR applications offer (such as *Alt Space*[1]). Another example is to be able to collaboratively build a structure, and explore and manipulate it in real-time in 3D (like with *Tilt Brush*[2] and *Oculus Medium*[3]). In order to effectively complete tasks via negotiation and collaboration, a significant level of trust is necessary between users. In this paper, we are interested in how different avatar representation can have an impact on user experience. By exploring how different configurations of avatar representations between paired users impact social interaction, we hope to bring valuable insight on establishing effective setups of avatar representation in CVE. In particular, we are interested in two aspects: self-presentation (whether to render a Self Avatar or not) and consistency (whether to maintain the same setup of self-representation between users in the CVE or not).

---

[1]https://altvr.com/
[2]https://www.tiltbrush.com/
[3]https://www.oculus.com/medium/

Previous research suggested that the use of Self Avatar can be a powerful tool in facilitating trust. Pan and Steed investigated the impact of the Self Avatar on collaborative and competitive tasks (Pan and Steed, 2017) in an immersive Virtual Reality (VR) system, and found that both the Self Avatar condition and the face-to-face condition led to higher trust scores than the no Self Avatar condition. A similar study in Augmented Reality (AR) investigating the effects of avatar representation on *Social Presence* - the feeling of being in VR with someone else - found that a realistic full body avatar was perceived as being the best for remote collaboration, but an upper body or artistic cartoon style could be considered as a substitute depending on the collaboration context (Yoon et al., 2019). However, in these studies, each dyad (pairing) had consistent self-representations. In this paper we explore the impact that consistency may have on trust within a collaborative setting. We ask whether consistency in self-representation could improve trust as well as the productivity between pairs in a CVE.

Another dynamic we examine is how results may vary when using a confederate (Study 1) as compared to paired participants (Study 2). The purpose of Study 1 was to validate the virtual test-bed and see whether the theory of consistency could be tested using a confederate, as it is a common practice in analysing participant responses to others due to the ease of preparation and recruitment. In Study 2, we developed the experimental design to include paired participants, giving us the opportunity to see whether organic social dynamics also matter when observing trust between a group. To investigate this we developed a CVE where two players can meet and play a collaborative game. Each participant will have either a self-representation of an avatar or virtual controllers in a consistent or inconsistent condition (see **Figure 1**). They will wear a Head-Mounted Display (HMD) which will allow them to see each other in VR and paired controllers to interact with the virtual objects in VR.

In summary, the main contribution of this work is that we evaluated the impact of the Self-Avatar on Trust, testing primarily for the effect of inconsistent and consistent representation between user pairs, which has not been done in the social VR setting. This investigation was explored across two studies to test for the effect of using a human confederate (Study 1) and participant-only pairs (Study 2). In this article we present the experimental design, analysis, and results of these studies. This includes testing two variants of a social dilemma task to objectively measure trust, and a personality trait questionnaire to find potential correlations between participant personalities and their interactions within VR.

2. BACKGROUND

2.1. The Self Avatar in CVE

We have reached a point in time where the use of VR has become a common medium for social interaction, increasingly so due to the 2020 COVID-19 pandemic which left most of the populace in isolation for increased periods of time. With the rise of consumer-ready and accessible HMDs (such as the Oculus Quest, Rift S, HTC Vive, and PlayStation VR), socialising and collaborating in CVE is becoming common practice with many international conferences transitioning to online formats. During these digital conferences, users are able to explore, chat and collaborate together in different tasks that require the need for clear and optimal implementations of avatar-mediated communication (see **Figure 2**).

In these virtual spaces there are varied displays of the Self Avatar. A good example to demonstrate this variation is *Alt space*[4]. In this CVE users can be represented by full body humanoid avatars, or robots with hands (but no arms). In other CVE, such as *RecRoom*[5], Self Avatars are depicted only from the torso up without arms. There are numerous research studies that investigate the impact the Self Avatar can have physically and psychologically on users. It was Botvinick and Cohen (1992) who first gave evidence that the phenomena of Body Ownership can be transposed onto other associated objects (i.e., the rubber arm). This effect was then explored with respect to 3D worlds. Yee and Bailenson (2007) examined how the alteration of the Self Avatar may influence the participants' self perception and behaviour towards others. They describe the phenomenon as the "Proteus Effect." In their first study, results revealed that participants assigned to more attractive avatars shared more intimately when engaging in self-disclosure with a confederate, whereas in the second study their results suggested that participants assigned taller avatars behaved more confidently in a negotiation task. The authors' conclude that participants' virtual representation was able to change how they interacted with each other within the CVE. From these studies, we can see evidence of how the setup of self-representation can have a strong psychological effect on social dynamics.

Other practical impacts of Self Avatars have been extensively investigated as well. In a further study (Kilteni et al., 2013), results showed that Caucasian participants who were embodied in a "casual dark-skinned avatar" had significant increases to drumming patterns in comparison with their baseline drumming. This could suggest having a different visual representation of the participants' identity in virtual reality may produce measurable beneficial outcomes within a CVE. More recently a study showed that an active Self Avatar which enables the use of gestures could alleviate cognitive load of a task and therefore improve performance (Steed et al., 2016). Mohler et al. (2010) showed that an animated Self Avatar was superior to that of a static one when participants took part in a task with distance estimation.

Studies have shown that the display of the Self Avatar can have a strong impact on social dynamics in CVE. For example, Steed et al. (1999) found that participants immersed in CVE tended to emerge in a leadership role when compared to those connected by desktop, showing how differences in self-representation could shift power dynamics.

In these studies participants were embodied as a full body Self Avatar, and we can reasonably suggest that this full body Self Avatar generates a positive outcome to the sense of *Presence*, interaction tasks and perceptual judgement. The exact definition of Presence is a matter of some debate, but
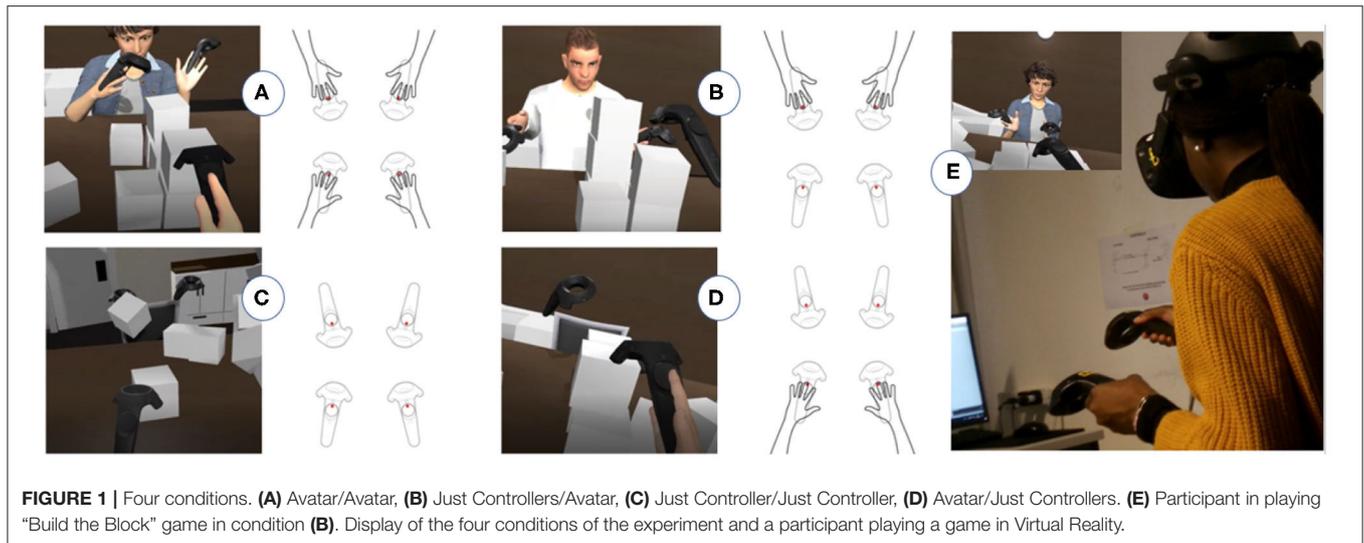
---

[4]https://altvr.com/
[5]https://recroom.com/

FIGURE 1 | Four conditions. **(A)** Avatar/Avatar, **(B)** Just Controllers/Avatar, **(C)** Just Controller/Just Controller, **(D)** Avatar/Just Controllers. **(E)** Participant in playing "Build the Block" game in condition **(B)**. Display of the four conditions of the experiment and a participant playing a game in Virtual Reality.
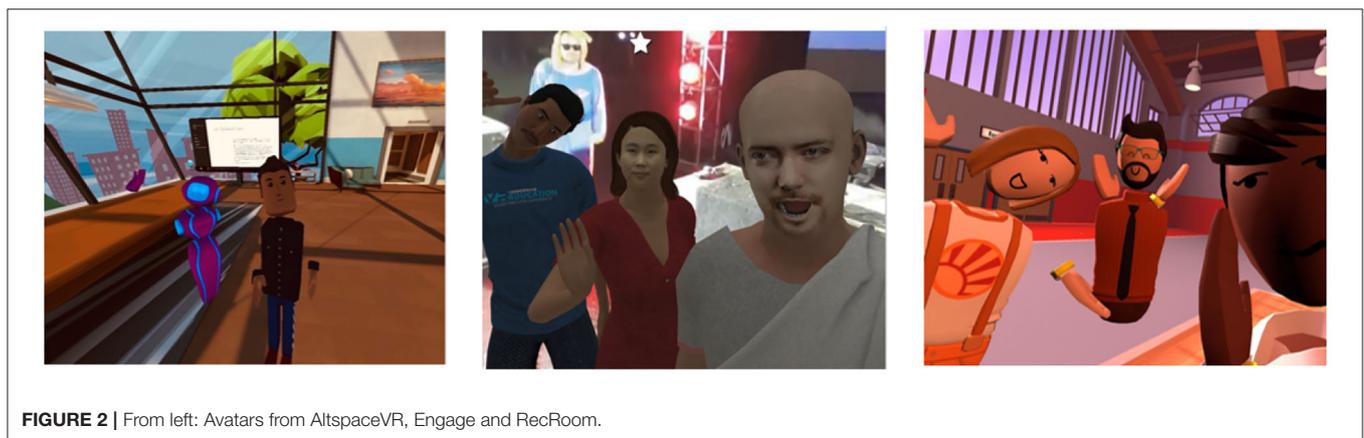

FIGURE 2 | From left: Avatars from AltspaceVR, Engage and RecRoom.

we follow Witmer's suggestion that Presence is the subjective feeling of being "present" in one environment (in this case VR), even when physically located in another (Witmer and Singer, 1998). Research suggests that participants may experience a sense of Body Ownership—the subjective feeling of ownership a participant may have over this virtual avatar (Kilteni et al., 2012) and agency—the motor control over the virtual body, over their virtual actions even in the total absence of visible virtual body-parts (Murphy, 2017). Additionally, the use of only visual hands and feet were sufficient to induce illusory Body Ownership, and this effect was observed as being just as strong as using a whole-body avatar (Kondo et al., 2018). From these studies we conclude that the visual representation and the setup of immersion can have a psychological impact on how participants complete tasks. In our study we chose to use full-body, gender-matched 3D models (who are holding controllers), as users' Self Avatars. Avatar rendering was turned off during our control condition, leaving only a pair of disembodied controllers as sole representation of the user. This allowed us to moderate how the participants would interact with the environment, and help isolate the impact that each representation condition may have.

We chose to investigate the impact "consistency" may have on a social interaction with collaborative tasks as it is a likely scenario to be affected.

2.2. Confederate vs. Participant

The use of confederates is a common occurrence in VR psychology studies, even though there is debate on how this may hinder a study's re-productivity (Doyen et al., 2012), or if participants may behave differently with confederates than another participant. Early research (Martin, 1970) suggests the possibility that the use of confederates to manipulate independent variables in small group experiments is compromised if the confederates arouse suspicion, and implies "deceived" and "undeceived" subjects do not behave alike. Though this predates the establishment and use of CVE, it probes at whether these social dynamics can carry over into a virtual space. It is still a common practise to utilise confederates in CVE studies, however research shows there are certain contexts in which their use could introduce unknown factors into data, such as when taking up the addressee role [if they know more than is warranted by the experimental task

and if their non-verbal behavior is uncontrolled (Kuhlen and Brennan, 2013)]. On account of this, we hypothesise that the use of a Self Avatar can become a potential hindrance when a confederate is utilised. We look deeper into this investigation by running two studies, one that used a confederate and the other paired participants.

2.3. Consistency in CVE

Consistency in avatar representation has been a topic of research for many years in VR. A branch of this research focuses on whether consistency in representation can have an impact on trust. Presently there is no consensus on this research question as there are various studies that depict results favouring either end of the debate. For example, Gong looked at how trust and judgement could be affected by avatar representation (Gong and Nass, 2007). Here the consistency was pairing a human face with a human voice or a humanoid (artificial human) face with a humanoid voice, and the inconsistency was the async of both these conditions. What they found was that in the inconsistent conditions, making judgement of the agent took a longer processing time and the participants felt less trust towards them. Here we see evidence in favour of this hypothesis. On the other hand, Latoschik compared paired interactions with abstract avatar representations based on a wooden mannequin, with high fidelity avatars generated from photogrammetry 3D scan methods. Participants were assigned one or the other and alternated between different representations of the virtual agent in dyadic social encounters (Latoschik et al., 2017). This created a 2 × 2 factorial design where two conditions were consistent (both participant and virtual agent had the same avatar) or inconsistent (participant and agent had different avatars). An interesting result found that the appearance of the virtual agent's avatar had an impact on the self-perception of the participant's own virtual representation. A more realistic-looking agent avatar seemed to increase the impressions of the changed Self Avatar and therefore helped to increase the suspension of disbelief for the respective avatar owners. However, they did not find any significant result regarding trust between the conditions. There is also research that revealed equal trust levels towards both categories of human and robot avatars. Nevertheless, participants still felt a significant sense of "togetherness" with the human-like avatar compared to the robot even though the participant only could see their human hands and the confederate had a full body (George et al., 2018). This condition in another perspective could still be considered inconsistent, however the fact that the participant could still see their hands may have played a role in the positive result (Kondo et al., 2018).

Additionally research has also shown that how the virtual environment is setup can affect the emotional states of participants (Dey et al., 2017) as well as the choice of the task that is to be completed (Regenbrecht et al., 2006; Kim et al., 2012). Fundamentally, it is important to be able to successfully immerse the participant into both the virtual world and the scenario of the scene, the susceptibility to these effects are mediated by two illusions of *Presence*. The two illusions are: *Place Illusion* and *Plausibility Illusion* (Slater, 2009). *Place Illusion* is heavily influenced by the technical setup of the virtual environment

or how participants are immersed. This involves fundamentally occluding the participants field of view (FOV) of the real world so that they are fully immersed in another location painted by the 3D graphical interface. This is commonly accomplished using a stereo Head Mounted Display (HMD) with a wide field of view. *Plausibility Illusion* centres around making the participant suspend their disbelief and feel that the scenario they are experiencing is a real event. Care must be taken that the virtual environment built is set up to provoke and support the intended need for the application. In Latokchik et al.'s experiment they used only basic non-verbal communication (hand wave) as the form of interaction, but there are other more complex scenarios that could be considered. Different communication scenarios may have an influence on online trust (Feng et al., 2004).

2.4. Measuring Trust

Trust is difficult to measure as it is a subjective construct. There are many different approaches to evaluating the development of trust, both objective and subjective. The most commonly used method to collect subjective data is questionnaires giving self-reports, but this is still a method under constant controversy on its validity. For example, Bailenson gave evidence in his research that objective measures, such as behavioural data (heart-rate, average movement) could be sensitive enough to pick up on responses that self-reports could not (Bailenson et al., 2005).

Behavioural tasks are another method used to gather objective data. Hale created a virtual maze as a behavioural tool for measuring trust. They manipulated virtual characters' trustworthiness during an interview stage with the participant and then measured how often they approached and followed advice from each character (Hale, 2017). In this study they compared their behavioural tool with using a Social Dilemma game called "The Investment Game." The investor was given 10 US dollars (different amounts have been used in subsequent studies, Johnson and Mislin, 2011) and had to decide how much of their 10 US dollars to send to the trustee, knowing that the amount they sent would be tripled before it was given to the trustee. Then, the trustee had to decide how much of the tripled amount to return to the investor. The game measures trust behaviour in terms of the percentage of money the investor is willing to send to the trustee. They found that where the maze picked up on specific trust, the investment game picked on trust felt by participants.

Due to the results above, we have decided to use both subjective and objective methods to collect data in this experiment with the intent to make our findings more robust. In the subjective questionnaire we also focus on the Liked score as research suggested that people who are liked more by others are also more likely to win their trust (Feng et al., 2004).

We are motivated by the research questions above to investigate the potential impact of self-representation and consistency in collaborative virtual environments, paying close attention to the effect on Trust and Liked as well as the possible influence on collaborative productivity between participants.

TABLE 1 | Table of conditions.

Self-representation	Consistent	Inconsistent
Avatar	Self Avatar + Other Avatar	Self Avatar + Just Controllers
Just Controllers	Just Controllers + Just Controllers	Just Controllers + Other Avatar

The '+' represents the confederate condition.

3. STUDY 1: CVE WITH A CONFEDERATE

In this study we looked at the effect of two factors on CVE: the self-representation of an avatar, and the consistency of representation between players. The participant was asked to play a collaborative game in VR with a confederate followed by a trust exercise, before completing a questionnaire. The confederate was a female researcher who was briefed on the task of the experiment and told to appear as another player. Our goal was to investigate how different configurations of avatar representations between dyads impact social interaction. Another addition to this study which extends the previous method is the use of a confederate. It would be interesting to see how Trust can be measured in this context as the result may inform future experimental designs in this area of research. Our hypothesis for Study 1 is as follows:

H1 Participants in consistent conditions (C1 and C3, see **Table 1**) will have a higher level of *Place Illusion*, *Plausibility*, and *Co-Presence*.
H2 Participants in consistent conditions will feel more trust towards the confederate, reflected in both their subjective trust score and offering a higher amount of money in the trust game, and will in turn report more positive feelings towards the confederate (measured by *Liked*).
H3 Participants in consistent conditions will perform faster in the collaborative game.

There is evidence to suggest embodying an avatar has a positive impact on subjective experiences, such as *Presence* (Skarbez et al., 2017), however we propose that even when participants do not have an avatar—but are in a consistent condition—they will feel higher *Presence* than in inconsistent conditions. Previous research by Slater has been done on aspects which can prevent or interrupt the flow of *Presence* (Slater and Steed, 2000). It could be that inconsistency in avatar representation between pairs could act as a "break in *Presence*," causing both the loss in plausibility and the feeling of being present in the virtual environment. If avatar realism is believed to be a balance of visuals and behaviour (Oh et al., 2018), then if there is a mismatch of expectation on both sides this may create a negative impact. We know from previous studies that there may be a correlation between how much a person is liked and trusted (Feng et al., 2004). We also know that successful non-verbal communication is positively impacted by the use of gestures as it helps to reduce cognitive load and makes conceptualising ideas easier (Steed et al., 2016). However, it may be that consistent avatar representation allows for mutually shared social cues which can be grasped and understood more

quickly, and therefore impacting play faster. In both conditions the controller is available to suggest hand orientation. Following on from **H3** we can further argue that with better communication established, participants may be able to act more efficiently. Research has shown that the Self Avatar can positively impact the experience of interacting with the virtual world (Steed et al., 1999; Kilteni et al., 2013), so we anticipate that, regardless of condition, the Self Avatar will have the stronger effect overall on trust.

3.1. Participants and Materials

A total of 17 participants were recruited for this experiment from George Mason University in North Virginia. Among them were six females and 11 males, with mean age 27 ±6.1. We had a female human confederate who the participant interacted with in each session in social VR. All participants were unacquainted with the confederate before the study, and never interacted with each other except via their particular experimental condition. Participants were not allowed to exchange social information either before or during the game. This study was approved by the ethics board of George Mason University. Participants were assigned a condition upon scheduling to make sure the breakdown in each condition was as even as possible. The order of the participants depended on their availability. The HTC Vive HMD was used to capture head position and rotation, due to lack of internal sensors, no gaze tracking was recorded. Additionally, position and rotation tracking data from the HTC Vive controllers was utilized to control the arm movements of the avatar allowing for 6DOF. The avatar (MORPH3D) was downloaded from the Unity Asset Store. Using high fidelity models have been seen to provoke more acceptance, especially if they are perceived as attractive (Latoschik et al., 2017). These models were given small face masks to limit this effect, as well as hide the (static) mouth from view. There was no eye movement implemented on the models. The virtual environment and game were created using Unity 3D version 2017.2.0f3. To enable the tracking data from the HTC Vive, we used the SteamVR Unity plugin. To allow for the 1:1 mapping of arm movement, we used the Inverse-Kinematics plugin InstantEdgeVR (now depreciated). This allowed for instantiation of first and third person avatars, synchronization of avatar movements across the network, movement prediction and synchronization of grabbing and letting go objects. There was no finger tracking, however the avatar's hand was manually connected to a controller—provided by the Steam plugin, to remove the affordance for movement. Tracking was done using the HTC Vive controller and headset which is over 60 Hz. Networking was provided by the Photon Unity plugin.

The experiment was held in the lab office of George Mason's Virginia Serious Games Institute (VSGI). The two users were placed facing each other in the center of the room. The HMDs were connected to two separate VR ready desktops at opposite ends of the room. The HTC Vive HMDs shared the same lighthouse sensors but were set up in SteamVR to face the direction of the center of the room, respectively.

3.2. Experimental Design

As shown in **Figure 1** and **Table 1**, the experiment had a between-subject 2 × 2 factorial design with each participant taking part in only one of the four conditions. The two factors are: *Self-Representation* (Self Avatar/Just Controllers) and *Consistency* (Consistent/Inconsistent). Participants either had a high fidelity, gender matched Self Avatar holding controllers or *Just Controllers* without an avatar. The confederate they interacted with either had the same consistent setting, or an inconsistent one. We did not manipulate the perceived representation and kept it consistent across conditions for both studies (i.e., if user A perceived themselves to have a Self Avatar, user B also perceived A to have a Self Avatar and vice versa). In the following we refer to having an avatar as AV, Just Controllers as JC.

3.2.1. Collaborative Game

Using the collaborative framework provided by Unity3D we created a short gaming experience for the experiment. The game, "Build the Block" was designed to be simple and enjoyable, with a timer included to add an element of game challenge. The participant would appear seated at a table with a confederate and be shown a series of sequences, which they would have to imitate with the blocks provided to them on the table. There were 10 possible models to replicate but we were only concerned with the first three sequences for data collection. This is because not all groups could finish within the time limit, but had enough time to manage at least three sets and experience a range of difficulty. The players could pick up and place the blocks in stacks using the VIVE controller where they would either see a virtual body with controllers or just a pair of controllers in the environment depending on the conditions. According to Myerson's publication "On the Value of Game Theory in Social Science," interactive tasks can be described where one player's action directly influences the others (Myerson, 1992). The mechanics of the game encouraged the participants to verbally and physically collaborate with each other in order to make progress. Participants were able to speak to each other through the mic in the headset and in all setups of immersion the controllers were visible therefore allowing a threshold of non-verbal communication consistent in all conditions. With this setup we hoped to highlight the effect of having consistent and inconsistent avatar representation between pairs whilst playing a collaborative game. This game was pilot tested in real life using Jenga blocks with two participants. They were asked to play with a confederate and then filled out a questionnaire on gamer experience (Ijsselsteijn et al., 2013) and their feelings towards the other player, including questions, such as *I thought it was fun, I thought it was hard*, and *I was good at it*. These were rated on a Likert scale from 1 to 5, 5 being fully agreed. The results were used to validate the choice of game and its design. Overall, they found the game enjoyable, engaging, but not really challenging. In response we made the block models slightly more complex, and grow in complexity as they are completed. Based on existing literature (Pan and Steed, 2017), we hypothesise that in consistent conditions (conditions 1 and 3), participants will be faster when both players have an Avatar.

3.2.2. The Investment Game

The experiment had two phases. The participants were first asked to complete a game of Build the Block with a confederate. They had to work together to lift the cubes and stack them on-top of each other, mimicking the sequence shown to them. The participant, depending on the condition, would have a different immersion setup which was either consistent or inconsistent with the confederate player.

To analyze the level of trust the participant feels towards the confederate, once the player finishes the game they will take part in a exercise called "The Investment Game."

The participant is rewarded with 100 points. They will be offered a chance to share some, all or nothing of this amount with the confederate (other player). Every time points are sent to another player it is doubled by the experimenter. The confederate (other player) will then be given the same option. The participant will have to assess the risk of losing their points in sharing with the other player. This was what was told to the participant. The amount the participant gives will be recorded. The participant will not be able to speak or see the confederate during this exercise. The goal was to get as many points as possible however there was no real world gain related to this exercise.

Example:

1. A decides to share 20 of the 100 points with B (A = 80, B = 20).
2. This is doubled and given to B (A = 80, B = 40).
3. B can then send back an amount of their points to A. B sends 20 (A = 80, B = 20).
4. This is doubled and given to A (A = 120, B = 20).

In this scenario B can also choose to keep all the points given, e.g., (A = 80, B = 40) If B did that, then A will end up with less points than what they started with and the more they give, the more they would lose out on their final point (note that there was no real world gain and the money was not subtracted from the participants' pay). This exercise tests the amount of trust A has in B to share their points. This was developed based on the Investment Game in Hale's study (Hale, 2017) and Glaeser's Trust Game (Glaeser et al., 2000). There was only one turn to share and potentially increase the initial amount of the participant. We observed the amount participants decided to share with the confederate, as a representative of the amount of Trust felt towards them. The more money given, the more the participant "trusts" that the other will reciprocate so that they both benefit.

3.3. Procedure

First, participants were given a brief and asked to sign a consent form and short questionnaire to collect demographic information. Participants were then informed that they would take part in a game in which they would have to stack blocks, according to the sequence shown to them for an undetermined period of time while seated. Once finished, they would take part in the investment game. After completing the game, participants were asked to fill out a questionnaire survey which gathered subjective data on their experience. At the end of the experiment, participants were paid for their time and debriefed (if desired). The whole process took ~30 min.

3.4. Measurements and Data Analysis

The level of trust was measured with both Questionnaire data (*Subjective Trust*) and also behaviour data (*Trust Money*) collected in "The Investment Game," as described in section 3.2.2. We also measured the extent to which participants *Liked* the other person (in this case, always the confederate), with a questionnaire (Pan et al., 2015). We also collected participants performance data in the Collaborative Game in VR (three sets), and other related questionnaire data [*Place Illusion*, *Plausibility* (Slater, 2009), and *Co-Presence* (Bailenson et al., 2005)]. All data analysis was performed with IBM SPSS Statistics version 23. We first conducted a two-way ANOVA test regardless of the normality of data distribution because ANOVAs are considered to be fairly "robust" to deviations from normality (see Maxwell et al., 2004 for a review), although no specific research has been conducted into the two-way ANOVA. In the instances where there has been a significant difference found in the data which were not normally distributed, we also ran a non-parametric test (Mann-Whitney U) for further analysis to validate the result.

3.5. Results—Behavioural

3.5.1. Investment Money

A two-way ANOVA was conducted on the *Investment Money* given, with the two between-group factors (*Consistency* and *Self-Representation*). There was no statistically significant effect for *Self-Representation* [$F_{(1,13)} = 0.97, p = 0.343, \eta^2 = 0.07$], nor for *Consistency* [$F_{(1,13)} = 0.61, p = 0.45, \eta^2 = 0.05$]. However, there was a significant **interaction effect** [$F_{(1,13)} = 9.22, p = 0.01, \eta^2 = 0.42$], suggesting a "confederate avatar effect": when the confederate did not have an avatar, more money was shared by the participant, indicating higher levels of trust (confederate with an avatar mean and standard error: 50.4 ± 8.8; confederate without an avatar: 86.3 ± 8.1).

The Shapiro-Wilk's test revealed that *Investment Money* was not normally distributed ($p = 0.001$). To verify the results from ANOVA we ran a two-tail Mann-Whitney U test on *Investment Money* between participants who interacted with a confederate with an avatar, and with those without. The result remained significant ($U = 12.5, p = 0.027$), confirming our findings from the ANOVA analysis.

3.5.2. Mean Game Time

A two-way ANOVA was conducted on *Mean Game Time* from collected timestamps from each of the three rounds played. There was no significant difference found in *Consistency* [$F_{(1,13)} = 0.00, p = 0.982, \eta^2 = 0.00$] and *Self-Representation* [$F_{(1,13)} = 0.37, p = 0.556, \eta^2 = 0.03$]. There was also no significance found for (*Consistency* × *Self-Representation*) [$F_{(1,13)} = 0.40, p = 0.536, \eta^2 = 0.03$]. We have also tested the game time of the three sets separately, and again no effect was found. We have also performed tests for each round, and no significant results was found (see **Figure 3** and **Supplementary Material** for a summary table).

3.6. Results—Questionnaire

3.6.1. Subjective Trust

A two-way ANOVA was conducted on *Subjective Trust* with the factors (*Consistency* and *Self-Representation*). No statistically significant difference was found for *Consistency* [$F_{(1, 13)} = 1.3, p = 0.27, \eta^2 = 0.09$]. However, for *Self-Representation*, there is some evidence indicating a "Self Avatar" effect [$F_{(1, 13)} = 4.4, p = 0.056, \eta^2 = 0.25$], suggesting that participants who had a Self Avatar were more likely to give a higher rating on trust to the confederate. There is also some evidence suggesting an interaction effect, indicating that the confederate gained more trust when without an avatar [$F_{(1, 13)} = 1.2, p = 0.07, \eta^2 = 0.23$]. Although not significant, these results are inline with our behavioural results from *Investment Money*. A Shapiro-Wilk's test reveals that data was not normally distributed ($p = 0.001$). We ran a Mann-Whitney U test on *Subjective Trust* to see if there was a difference in score between participants with a Self Avatar (AV) and without (JC). Though there was a higher *Subjective Trust* scores for AV (*meanrank* = 10.78) than JC (*meanrank* = 7.00), they were not statistically significantly different ($U = 20, p = 0.139$).

3.6.2. Liked

A two-way ANOVA was conducted on *Liked* from the questionnaires, with the factors (*Consistency* and *Self-Representation*). There was no statistically significant main effect of *Consistency* [$F_{(1, 13)} = 0.3, p = 0.579, \eta^2 = 0.02$]. However, there was a statistically significant main effect of **Self-Representation** [$F_{(1, 13)} = 1.0, p = 0.008, \eta^2 = 0.43$]. There was no interaction effect [$F_{(1, 13)} = 1.3, p = 0.269, \eta^2 = 0.09$]. Data was normally distributed as assessed by Shapiro-Wilk's test ($p > 0.5$).

The significant self-representation effect revealed here indicates that participants with an avatar (AV) were more likely than JC to give more positive marks to the confederate (AV: 6.6 ± 0.2; JC: 5.7 ± 0.2), regardless of the confederate having an avatar or not. This is inline with findings on *Subjective Trust* presented in section 3.6.1.

3.6.3. Place Illusion

A two-way ANOVA was conducted on *Place Illusion* from the questionnaires, with factors *Consistency* and *Self-Representation*. We found a significant difference in both **Consistency** [$F_{(1, 13)} = 5.0, p = 0.004, \eta^2 = 0.27$], and **Self-Representation** [$F_{(1, 13)} = 11.83, p = 0.036, \eta^2 = 0.48$]. No interaction effect was found [$F_{(1, 13)} = 0.07, p = 0.886, \eta^2 = 0.005$]. Data was normally distributed, as assessed by Shapiro-Wilk's test ($p > 0.05$).

As shown on the Boxplot (see **Figure 4**), participants in the consistent condition reported a higher level of *Presence* (consistent: 5.0 ± 0.3; inconsistent: 3.7 ± 0.3), supporting **H1**. Participants without an avatar (JC) seem to have reported a higher level of *Place Illusion* than AV (AV: 3.9 ± 0.3; JC: 4.8 ± 0.3).

3.6.4. Plausibility

A two-way ANOVA was conducted on *Plausibility* with *Consistency* and *Self-Representation*. No statistically significant was found for *Consistency* [$F_{(1, 13)} = 0.43, p = 0.525, \eta^2 = 0.32$] or *Self-Representation* [$F_{(1, 13)} = 2.09, p = 0.17, \eta^2 = 0.14$]. Neither was there an interaction effect [$F_{(1, 13)} = 0.19, p = 0.892, \eta^2 = 0.001$].
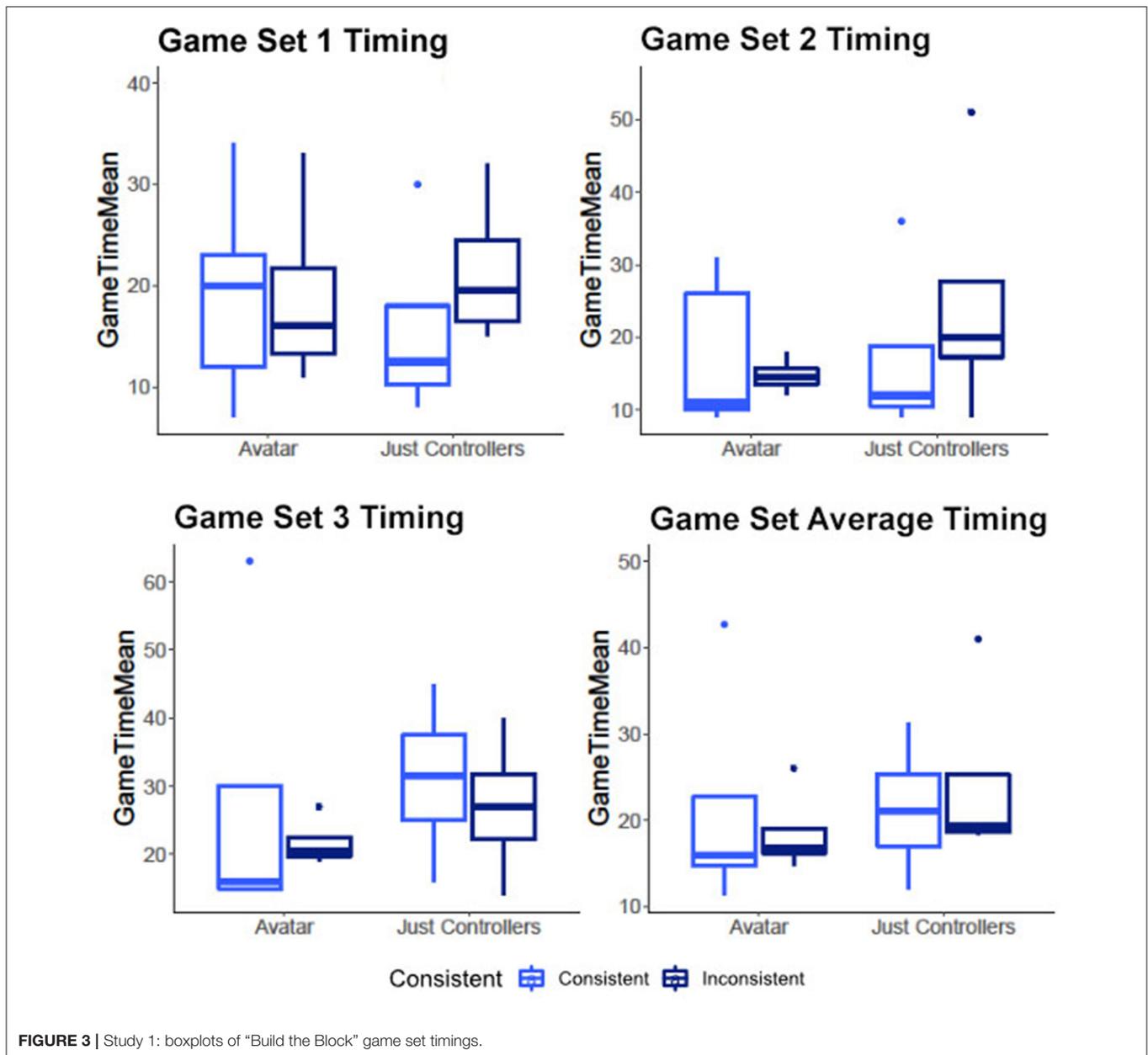
FIGURE 3 | Study 1: boxplots of "Build the Block" game set timings.

3.6.5. Co-presence

A two-way ANOVA was conducted on *Co-Presence* with *Consistency* and *Self-Representation*. No significant effect was found for *Consistency* [$F_{(1, 13)} = 0.274, p = 0.09, \eta^2 = 0.91$], or *Self-Representation* [$F_{(1, 13)} = 2.93, p = 0.111, \eta^2 = 0.18$]. Neither was there an interaction effect [$F_{(1, 13)} = 1.30, p = 0.274, \eta^2 = 0.91$].

3.7. Discussion

The results from the questionnaire revealed a significant difference in mean *Liked* score when participants have an avatar (AV compared to JC), regardless of the condition of the confederate. Though not significant this pattern is almost

mirrored in the results for mean *Subjective Trust* score. This could be due to the fact that the participants were able to express themselves through non-verbal cues, such as gesturing or looking at the confederate. This could have resulted in the confederate being able to better coordinate with the participants and provide appropriate verbal and non-verbal feedback. There are many studies demonstrating the importance and effect of gesturing, an example being that the mimicry of gestures and body language could be an indicator of trust (Verberne et al., 2013). Another recent discovery is the potential ability to reduce cognitive load whilst completing a task (Steed et al., 2016). An alternative reasoning is that the confederate was perhaps able to respond to the participants gaze—suggested by the movement
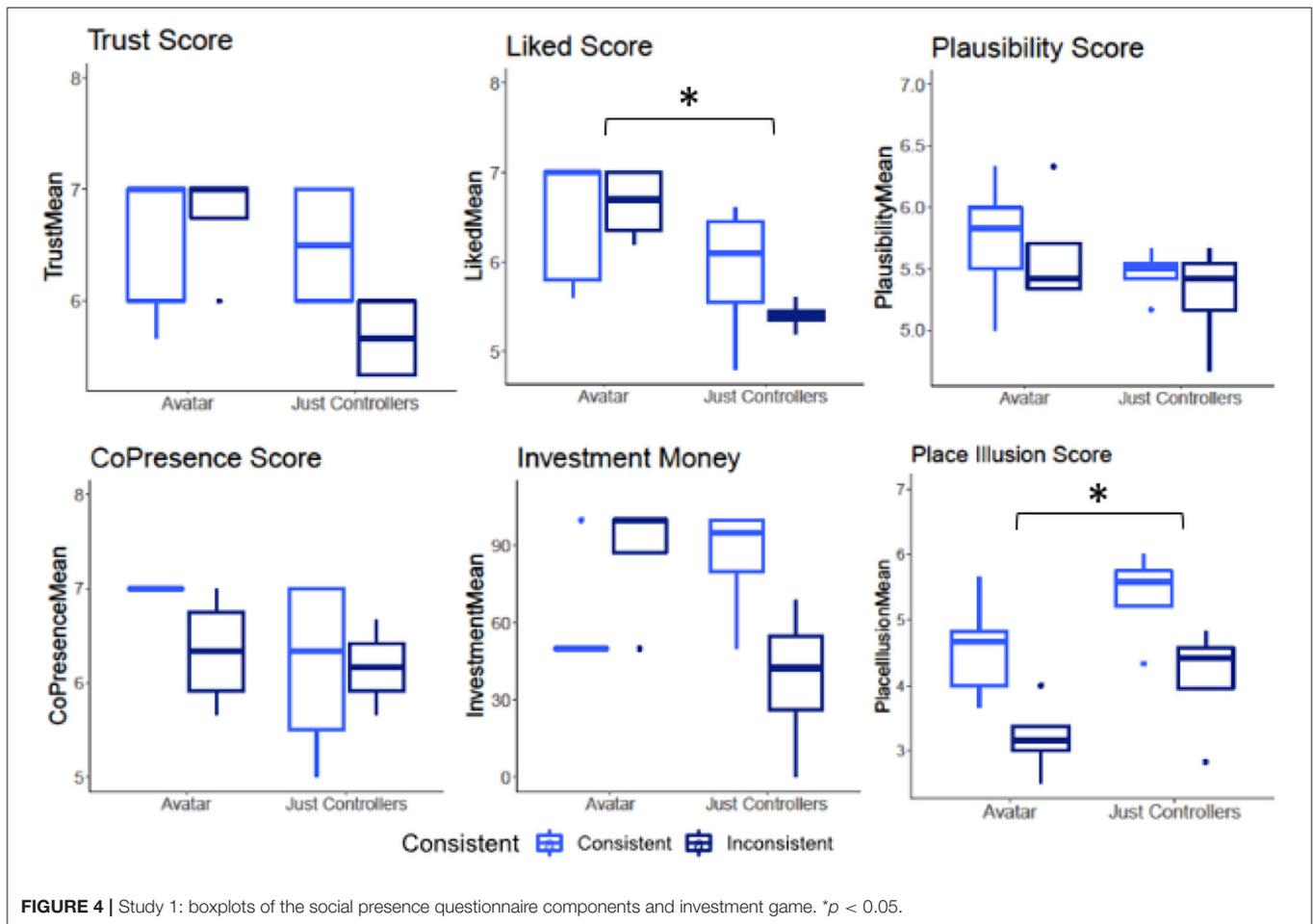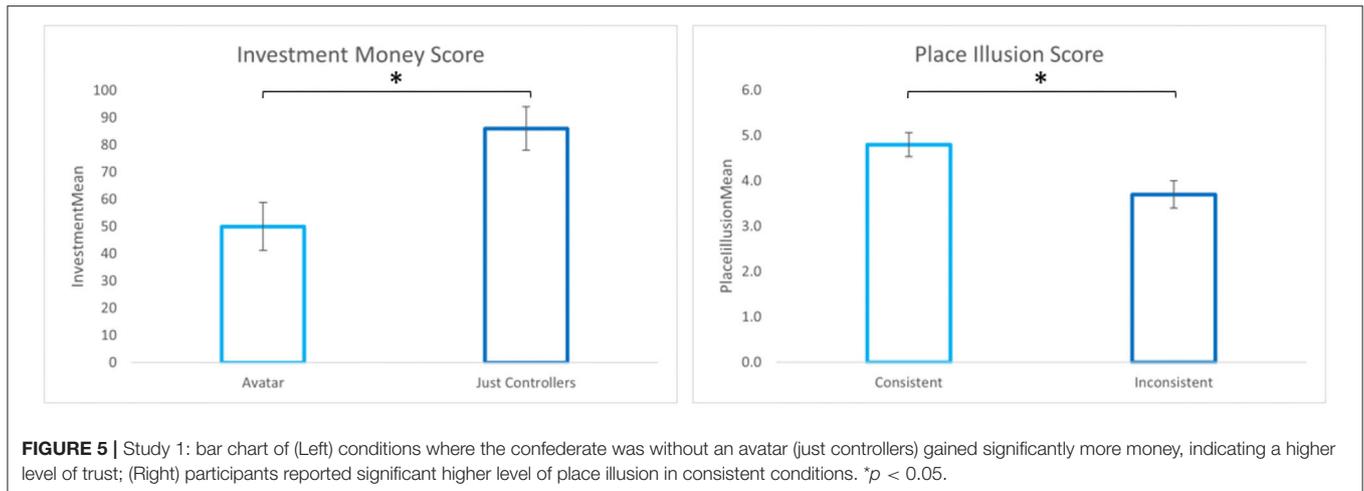
FIGURE 4 | Study 1: boxplots of the social presence questionnaire components and investment game. *$p < 0.05$.

and rotation of the avatars head in a more appropriate way. There has been many investigations on the positive impact of eye gaze on avatar-mediated communication (Garau et al., 2001). Participants with an avatar also reported feeling higher levels of *Plausibility*, *Co-Presence* but surprisingly not *Place Illusion*. However, as we believed, there were high scores across all three components with those who were in consistent conditions supporting **H1**. This is unexpected as previous studies have shown that a Self Avatar can positively impact *Place Illusion*. This could perhaps be explained by the potential cognitive load on participants, or an effect of the technical setup of the avatar representation.

There was a significant interaction effect between *Consistency × Self-Representation* on *Investment Money* (see **Figure 5**). When we observe the mean data from the Investment Game we can see that overall, more money was shared by the participant when the confederate did not have an avatar, suggesting that when the confederate had JC, they were better at gaining trust. The 3D models utilised in the experiment were from the high fidelity Morph3D package from the Unity asset store. Using high fidelity models have been seen to provoke more acceptance, especially if they are considered "attractive" (Latoschik et al.,

2017). These models were given small face masks to moderate for this effect as well as hide the non-animated mouth from view, which may have hindered trust. It is also possible that the reduce in trust was caused by the uncanny valley effect, where sometimes the use of more realistic models could potentially trigger eeriness (Masahiro, 1970). It could also be argued that this is due to the use of the confederate. In the setup of the study the confederate is instructed to play a game with each participant whilst pretending they are playing it for the first time. There is research to suggest that when a confederate is being deceitful this may provoke the participant to act differently and that "suspicious" confederate behaviour may be more likely compromise results (Martin, 1970). In this case this effect may have been heightened due to the confederate having a avatar. It could be that the deception overrode the impact of consistency. There is also research which suggests there is a risk of using confederates who are too familiar with the task (Kuhlen and Brennan, 2013).

These findings support the importance of research in exploring the impacts of a Self-Avatar, but also the results brings our attention to the use of a confederate and the nature of their representation, which may inform future research in this field.

FIGURE 5 | Study 1: bar chart of (Left) conditions where the confederate was without an avatar (just controllers) gained significantly more money, indicating a higher level of trust; (Right) participants reported significant higher level of place illusion in consistent conditions. *$p < 0.05$.

4. STUDY 2: CVE WITH PAIRED PARTICIPANTS

Following the completion of Study 1, we ran a main study with improvements to the testbed and experimental design. We hoped to both validate our initial findings and expand on the results by using paired participants. The rest of the changes we included are listed as follows:

1. We set up a collaborative game in virtual reality where participants would be run in pairs instead of using a confederate, giving us more data and removing the potential for confederate bias.
2. We removed the masks from the full body avatars.
3. We used the DayTrader game as means to objectively investigate trust, and ran this exercise three times during the session. The repetition gives more insight on the changes in trust through the experience, improving on our initial trust exercise process.

In each condition the players will either have a high fidelity, gender matched Self Avatar who will be holding controllers, or Just Controllers without an avatar. This will also be consistent or inconsistent between-subjects 2 × 2 factorial design. The aim for this experiment was to continue to investigate the impact of Self-Representation in paired consistent and inconsistent collaborative conditions. The hypothesis for this study is as follows:

H1 Paired participants in Consistent conditions will feel more trust towards each other. Those participants will therefore invest more in the DayTrader game due to increased trust, than those in inconsistent conditions.

H2 Participants in inconsistent conditions will feel less trust.

Though similar to those in Study 1, we wished to evaluate how the findings in Study 2 will differ with the use of paired participants.

4.1. Participants and Materials

A total of 18 participants took part in this experiment. All participants were recruited from Goldsmiths College, University of London. Among them were nine females and nine males. Ages ranged from 18 to 34 ($M = 25$, $SD = 5.0$). All pairs were unacquainted with each other before the study, and never interacted with each other except via their particular experimental condition. Participants were not allowed to exchange social information either before or during the game. The experiment was held in the Virtual and Augmented Reality Lab which consisted of two separate rooms. The paired participants were allocated with one individual in each room. The HTC Vives were connected to two separate virtual reality ready desktops at opposite ends of each room to maximise distance, and prevent any noise carrying through the walls. Participants were assigned a condition upon scheduling to make sure the breakdown in each condition was as even as possible. The order of the participants depended on their availability.

4.2. Experimental Design

Similar to Study 1, this experiment was a between participants 2 × 2 factorial design with the same factors (see **Table 1**). However, this time, instead of a confederate, each participant was paired with another participant. Another difference from Study 1 is that we replaced the Investment Game with the DayTrader game, following the 2017 study conducted by Pan and Steed (2017) which also used paired participants. This is because we wished to follow the method setup in Pan's work (Pan and Steed, 2017) in which this study is attempting to build upon.

4.2.1. DayTrader Game

The DayTrader game is a social dilemma task in which the short-term interests of individuals conflict with the long-term interests or goals of the group. We chose this social dilemma scenario because it provides measures of trust that have been tested for reliability and validity. The use of the DayTrader game was inspired by previous work (Johnson and Mislin, 2007; Rae et al., 2013) and used in a recent study (Pan and Steed, 2017). We decided to change the investment game from Study 1 to roughly follow the experimental design of this study in-order to extend the findings of Pan's work. The three staged method allowed us to see the changes in trust over time.

The game involved three sets of five rounds. For each set of the DayTrader game, each participant was given 30 credits that they could either keep or put into a pool that was shared between the two participants. At the end of the round, credits that they chose to keep doubled in value, while the credits in the shared pool tripled and were then split evenly between the two participants. At the end of each set of five rounds, the participant that earned the most credits in that set of five, received a 300 credit bonus. This bonus had the effect of giving an extra profit to the participant who contributed less than his or her partner. If both participants earned the same amount, they both received the bonus. Each participant is only told their new amount at the end of a round. They are not allowed to ask the other participant's amount or be given any means to work out the math.

Example:

1. A gives 20 of the 30 points to the shared pool. B gives 10. Shared pool total is now 30 (A = 10, B = 20).
2. Kept money is doubled (A = 20, B = 40).
3. Shared pool amount is tripled and shared between A and B (30 × 3/2, A = 20 + 45, B = 40 + 45).
4. New amounts for round 2 are A = 65, B = 85.

In this scenario B can also choose to keep all the points given and get more. However, both will gain more by giving all equally. This exercise tests the amount of trust A has of B and vice versa. Similar to Pan and Steed (2017), each pair of participants would play this game three times (i.e., three sets of five rounds). Participants will not be able to speak or see each other during this exercise except in the final (third) set, where they will be allowed a 30 s phone call mediated by the experimenters, as detailed in section 4.3 Procedure.

4.3. Procedure

Participants were led into different rooms by two separate researchers, with a gap of 5 min between them, to ensure they did not interact. They were briefed and asked to sign a consent form and fill a short questionnaire to collect demographic information, including the A 10-item Short Version of the Big Five Inventory spectrum (Rammstedt and John, 2007). Once complete, they were both given a sheet of paper explaining the rules of the DayTrader game. After confirming both participants understood the rules, the participants played five rounds (first set) of the DayTrader game with each other over voice-only communication, with each researcher recording the progress and results.

Participants were then given a sheet of paper explaining the "Build the Block" game and after confirming they understood the rules, were helped into the VR setup. They were given the opportunity to learn how to play the game with a "demo round." In this demonstration round, participants were asked to build a pre-existing shape completely alone (i.e., the other participant was not present during this time)—this demo round was not timed or included in analysis. Participants could not continue until they demonstrated an understanding of how to use the Vive controllers to pick up blocks, and how to use the Vive controllers to progress levels (change sequences) as part of the demo round

completion. Following this, the researchers prepared to run the main task. The participants were once again reminded of the instructions before they began. They were encouraged to speak to one another and strategize on how they would complete the task efficiently over the voice communication setup, as well as utilising the VR environment. They had 10 min to complete 10 levels of "Build the Block." Participants completed the task in either of the four conditions, while remaining seated for its entirety. Similar to Study 1, participants were allowed to speak to each other during the VR game, through the microphone provided by the HTC headsets. The HTC wands allowed them to move their arms which were either represented by a full body avatar with a controller, or just a controller. During this time, participants performance outside of VR was also recorded on video. Once finished, participants were asked to fill out a questionnaire survey which gathered subjective data on their level of *Presence*, *Subjective Trust*, and interaction towards each other when playing. Participants were then asked to play the second set of the DayTrader game.

After this was completed, participants were given the opportunity to communicate over voice for 30 s, in order to develop a strategy prior to starting the third and final set. In order to ensure participants had sufficient time to confer with one another, we tested the time-limit through improvised conversation prior to the study. Afterwards, most participants felt the 30 s was enough—which was reflected in fact that only one participant group reached the time-limit overall. After a consensus was reached between participants (or time ran out) they then played the final five rounds of the DayTrader game (third set). When evaluating the DayTrader exercise, it was our intent that the first set gives a baseline of trust; the second set establishes trust based on the VR encounter; and the third set validates this trust built in the second set. At the end of the experiment, participants took part in a semi-structured interview with the researchers, were paid for their time and debriefed if desired. The HTC Vive headsets were wiped with a cleaning cloth and other touched equipment with an antibacterial wipe after each participant. This session took roughly 45 min.

4.4. Measurements and Data Analysis

All measures and data analysis follow the same as Study 1, other than "The DayTrader game" (described in section 4.2.1) instead of the previous trust game.

4.5. Results—Behavioural Data

4.5.1. DayTrader Game Results: Investment Money

We only used the final round (round five) from each of the three sets to look at participants' level of trust: before the experiment (Set 1), after VR (Set 2), and finally after the phone call discussion (Set 3) as seen in **Figure 6**. A two-way ANOVA was conducted on the *Investment Money* for each set, with *Consistency × Self-Representation* as between-subjects factors.

For Set 1, no effect was found for *Self-Representation* [$F_{(1,14)} = 1.404, p = 0.256, \eta^2 = 0.091$], or *Consistent* [$F_{(1,14)} = 0.32, p = 0.582, \eta^2 = 0.02$], and no interaction effect was found [$F_{(1,14)} = 0.67, p = 0.426, \eta^2 = 0.05$]. As expected,
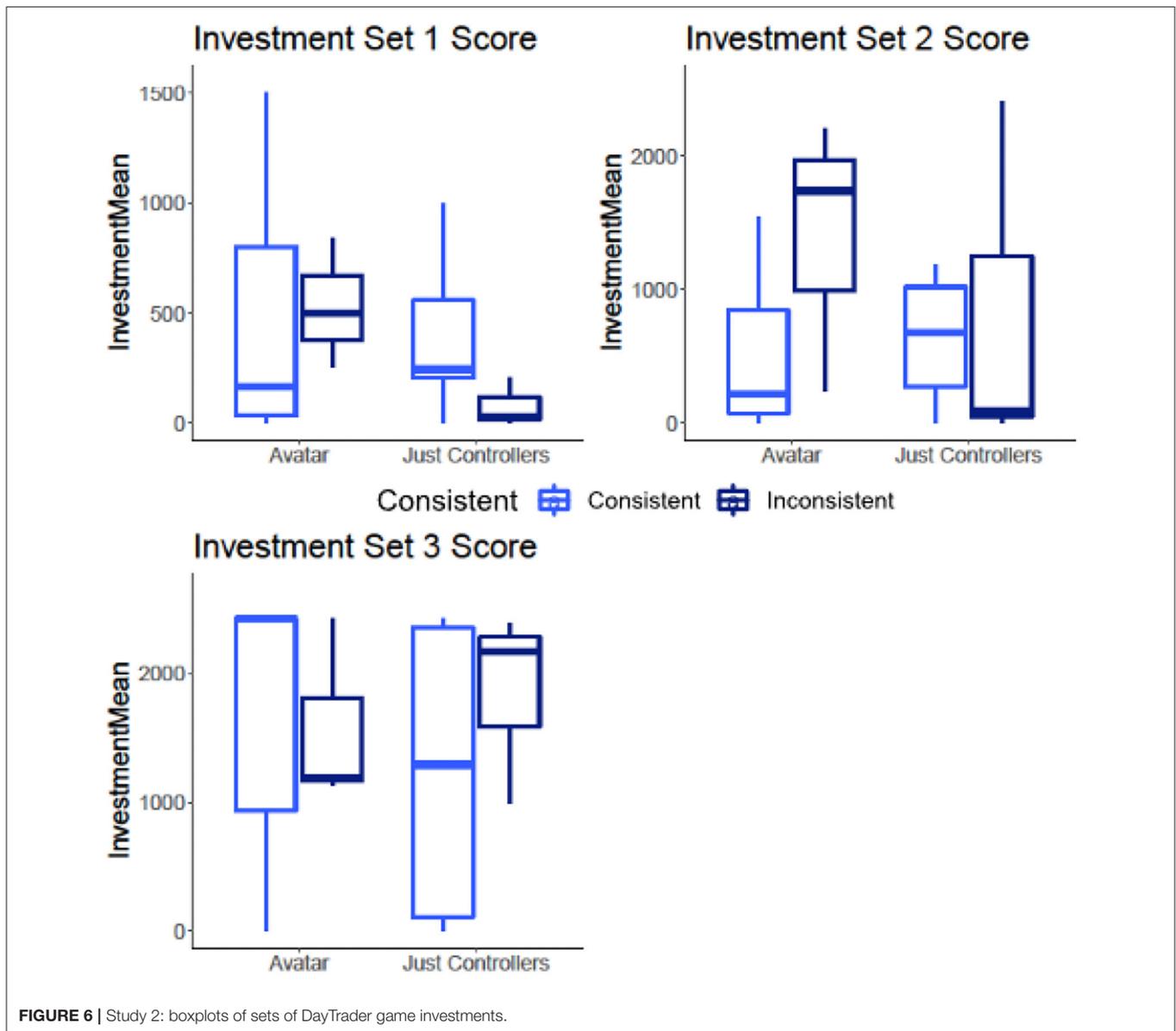
FIGURE 6 | Study 2: boxplots of sets of DayTrader game investments.

before the VR collaboration game interaction, there were no significant differences among the four conditions, with respect to trust between participants.

However, contrary to our hypothesis, for Set 2, no effect was found for *Self-Representation* [$F_{(1,14)} = 0.3, p = 0.595, \eta^2 = 0.021$] or *Consistent* [$F_{(1,14)} = 1.82, p = 0.198, \eta^2 = 0.12$], and no interaction effect was found [$F_{(1,14)} = 0.73, p = 0.406, \eta^2 = 0.05$].

Similarly, Set 3 reveals no effect *Self-Representation* [$F_{(1,14)} = 0.03, p = 0.66, \eta^2 = 0.01$], *Consistent* [$F_{(1,14)} = 0.23, p = 0.642, \eta^2 = 0.02$], or an interaction effect [$F_{(1,14)} = 0.45, p = 0.513, \eta^2 = 0.03$].

4.5.2. Mean Game Time Results
There were three sets assessed from the collaborative "Build the Block" game. Set 1, Set 2, and Set 3 as seen in **Figure 7**. A two-way

ANOVA was performed for each set on *Mean Game Time* with *Consistency × Self-Representation* as between-subjects factors.

For Set 1 we found no interaction effect [$F_{(1,14)} = 0.026, p = 0.874, \eta^2 = 0.002$], and no effect for *Self-Representation* [$F_{(1,14)} = 0.026, p = 0.874, \eta^2 = 0.002$]. However, there was a statistically significant effect on **Consistency** [$F_{(1,14)} = 6.21, p = 0.028, \eta^2 = 0.341$], suggesting that participants in the inconsistent conditions were able to complete their task faster (inconsistent: 39.0 ± 8.1, consistent: 64.7 ± 6.4). Data was normally distributed as assessed by Shapiro-Wilk's test ($p < 0.05$).

Similarly, for Set 2 there was no interaction effect [$F_{(1,14)} = 0.31, p = 0.863, \eta^2 = 0.01$], and no effect on *Self-Representation* [$F_{(1,14)} = 0.31, p = 0.863, \eta^2 = 0.01$]. However, there was an effect on **Consistency** [$F_{(1,14)} = 12.16, p = 0.004, \eta^2 = 0.47$], but this time participants in the inconsistent conditions were slower (inconsistent: 77.0 ± 10.9, consistent: 30.7 ± 7.7). Data was

FIGURE 7 | Study 2: boxplots of "Build the Block" game set timings.
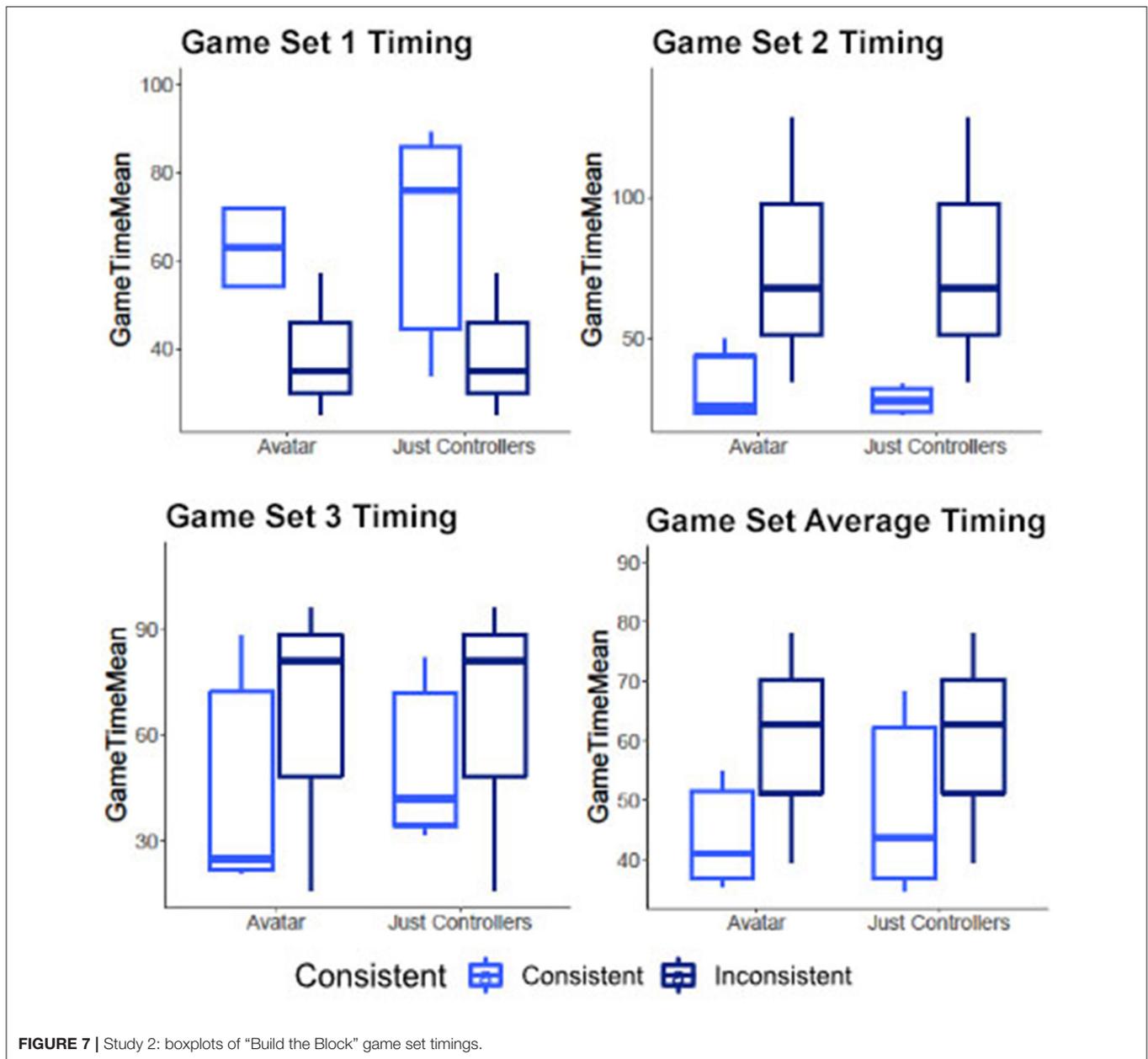
not normally distributed as assessed by Shapiro-Wilk's test ($p <$ 0.05). A Mann-Whitney U test on *Consistency* confirmed our finding ($U = 4, p = 0.003$, using an exact sampling distribution for U).

No interaction effect was found for Set 3 [*Self-Representation* × *Consistency*: $F_{(1,14)} = 0.48, p = 0.83, \eta^2 = 0.01$; *Self-Representation*: $F_{(1,14)} = 0.48, p = 0.83, \eta^2 = 0.01$; *Consistency*: $F_{(1,14)} = 0.91, p = 0.355, \eta^2 = 0.06$].

From **Figure 7** we can see that significant results of *Consistent* for Set 1 indicated that in consistent conditions, participants took longer to complete the set. This effect however reverted in Set 2 where participants performed faster in consistent conditions, before finally vanishing in Set 3.

We have also tested the game time average, and again no significant difference was found (See **Supplementary Material** for a summary table).

4.6. Results—Questionnaire
4.6.1. Subjective Trust
A two-way ANOVA was conducted on *Subjective Trust* with factors *Consistency* and *Self-Representation*. No interaction effect was found [$F_{(1,14)} = 0.1, p = 0.761, \eta^2 = 0.01$], and no effect was found for *Self-Representation* [$F_{(1,14)} = 0.1, p = 0.761, \eta^2 = 0.01$]. However, there was an effect of **Consistency** [$F_{(1,14)} = 9.62, p = 0.008, \eta^2 = 0.879$]. Data was normally distributed, as

assessed by Shapiro-Wilk's test ($p > 0.05$). This indicated that participants in consistent conditions reported a higher level of *Subjective Trust* (consistent: 6.2 ± 0.2, inconsistent: 5.1 ± 0.3), supporting our **H1**.

4.6.2. Liked

A two-way ANOVA was conducted on *Liked* with factors (*Consistency* and *Self-Representation*). No statistically significant effect was found for *Consistency* [$F_{(1,14)} = 0.02, p = 0.889, \eta^2 = 0.01$], *Self-Representation* [$F_{(1,14)} = 0.851, p = 0.372, \eta^2 = 0.057$], and *Consistency* × *Self-Representation* [$F_{(1,14)} = 0.01, p = 0.944, \eta^2 = 0.01$].

4.6.3. Place Illusion

A two-way ANOVA was conducted on *Presence* with *Consistency*, and *Self-Representation*. No statistically significant result was found. *Consistency*: $F_{(1,14)} = 0.001, p = 0.98, \eta^2 = 0.001$, *Self-Representation*: $F_{(1,14)} = 0.001, p = 0.98, \eta^2 = 0.001$, and *Consistency* × *Self-Representation*: $F_{(1,14)} = 1.13, p = 0.31, \eta^2 = 0.074$.

4.6.4. Plausibility

A two-way ANOVA was conducted on *Plausibility*. No effect was found for *Self-Representation* [$F_{(1,14)} = 0.01, p = 0.921, \eta^2 = 0.01$] nor *Consistency* [$F_{(1,14)} = 0.83, p = 0.379, \eta^2 = 0.06$]. However, there was an interaction effect found between (**Consistency** × **Self-Representation**) [$F_{(1,14)} = 5.4, p = 0.036, \eta^2 = 0.28$]. Data was normally distributed, as assessed by Shapiro-Wilk's test ($p > 0.05$).

As shown on **Figure 8**, this suggested that participants reported the experience to be more plausible when the person they interacted with was without an avatar.

4.6.5. Co-presence

A two way ANOVA was conducted on *Co-Presence* with the two factors. No statistically significant results were found over *Consistency* [$F_{(1,14)} = 0.45, p = 0.506, \eta^2 = 0.03$], *Self-Representation* [$F_{(1,14)} = 0.01, p = 0.924, \eta^2 = 0.01$] or *Consistency* × *Self-Representation* [$F_{(1,14)} = 0.37, p = 0.301, \eta^2 = 0.08$].

4.7. Big Five Personality Questionnaire

A Pearson's product-moment correlation was run to assess the relationship between the Big Five personality measure using the Ten-Item Personality Inventory (TIPI), and the Social Presence questionnaire components.

Preliminary analyses showed the relationship to be linear with both variables normally distributed, as assessed by Shapiro-Wilk's test ($p > 0.05$).

There was a statistically significant, moderate negative correlation between Liked score and Openness, [$r_{(16)} = -0.48, p < 0.05$]. This suggests, the more subjectively open the personality of the participant was, the less likely they were to like the other participant.

There were no other correlations found between the other variables (see **Figure 9**).

4.8. Interview Feedback

After the experiment, participants took part in a semi-structured interview to gather some additional feedback about their experience. In this paragraph we will explore some of the themes that arose through their answers. Responses were recorded by the researchers after the experiment and later coded into recurring themes. The high-level themes were as follows:

Participants felt that overall the first DayTrader game did not affect their interaction in VR. Most participants either "did not relate the two experiences," (p11) or felt that they "still didn't know what the person was like," (p13). This is important in validating that the first impression received of the other player was experienced through VR and whether or not they gained the bonus did not colour their interaction.

Participants felt that the VR session made the player seem more "real" and gave them an impression of the other player. The participants felt working together on a task made the person seem real. Some felt the "person was a blank slate before but started filling with detail," (p2) as they played. They were able to become "familiar with [their] personality and thinking," (p1).

Participants felt a shift from Competitive to Collaborative when playing the VR Game. Most participants started off with a competitive mentally with a goal to win. It is interesting to note that participants thought of the DayTrader game as a competitive activity as it could explain the variance in the results between conditions. One participant mentioned that their partner was "friendly in the VR version, more collaborative, and a team player. But in the DayTrader game [they] seemed a bit more calculated and logical," (p3). Some participants also suggested that they believed players acted differently or had different strategies in each separate game.

Participants' VR experience had the greater impact overall on their impression of the other person, but the phone call also helped in solidifying their feelings. Participants felt that over "just speaking," having an interaction with the other player helped them foster a sense of collaboration and made the other seem more real. The phone call was "reassurance" for many in their opinion. "It's hard to say, they both were effective in different ways. The VR gave me an impression of their actions, and then the 30 s phonecall was very informative—and then the follow-through on the phone call kinda cemented my opinion of them," (p18).

4.9. Discussion

In this study we observed four conditions of avatar representation between dyads: AVxAV, AVxJC, JCxJC, JCxAV. Participants were tasked to collaboratively complete a game in virtual reality in one of these four conditions, and their sense of trust were assessed both objectively through the DayTrader game and subjectively through use of questionnaires.

Surprisingly, contrary to our Study 1, we found no significant effects on how much participants were willing to cooperatively invest. This finding does not support H2. Hale proposed in her research that there are different kinds of trust that can be measured and perhaps therefore this method may not be robust enough to filter all types effectively (Hale, 2017). More research
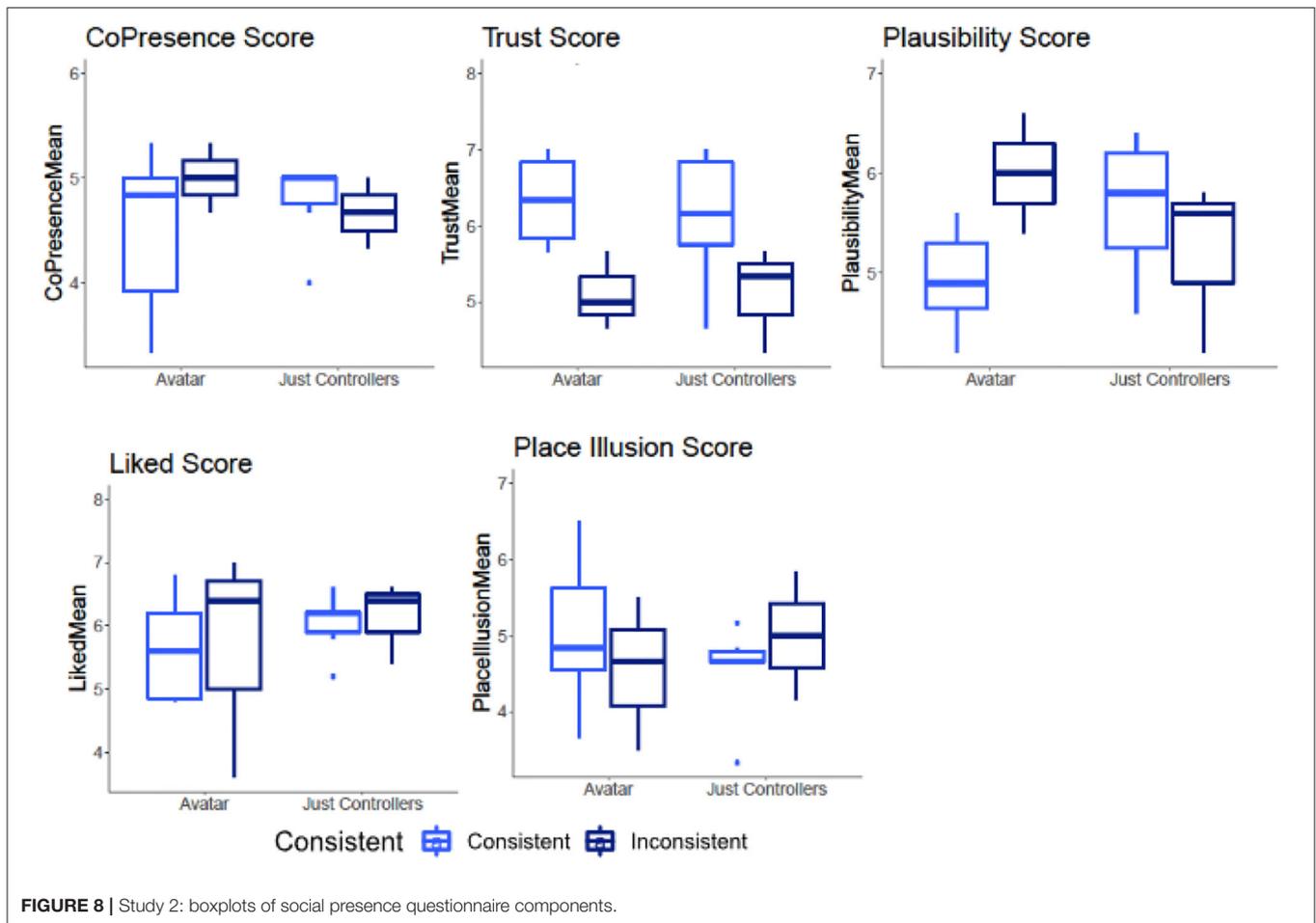
FIGURE 8 | Study 2: boxplots of social presence questionnaire components.

needs to be done using DayTrader as a valid metric for measuring trust in an avatar mediated virtual environment.

Our secondary behavioural measure was the time taken for each of the three sets of the "Build the Block" game. We can see that in Set 1 (see **Figure 10**), it was participants in the inconsistent conditions which were able to finish faster.

One possible explanation relates to Sadagic et al.'s work on leadership in CVE (Steed et al., 1999). They found that, in inconsistent CVE, the participant in the most immersive condition took a leadership role. It is possible that in our study the participant with the avatar naturally took on a leadership role. This would initially simplify the social dynamics in the unfamiliar game condition and enable the participants to work more quickly without the need for implicit negotiation of collaborative roles. On the other hand, the consistent participants may be putting more effort into establishing how to work together. More research would be needed to confirm whether this is the case.

This pattern however has been swapped in Set 2 where the participants in consistent conditions were significantly faster to finish their task. It's best to keep into consideration that as the rounds increase so does the complexity of the shapes to recreate. We see here that the initial advantage of inconsistent conditions disappears and consistent pairs are able to work

faster, presumably because they are able to work together more effectively once a pattern of interaction has been established in the first round.

In Set 3 there was no significant difference between groups on their time. However, we can see from the data (**Figure 7**) that participants in consistent conditions still continued to play faster than those in the inconsistent conditions, suggesting that overall, consistency has a positive effect on productivity in CVE.

The correlation analysis with the Ten-Item version of the Big Five Inventory showed a moderate negative correlation between the Liked Score and the self-report of Openness, which was unexpected. The results were significant ($p < 0.05$), suggesting the potential for deeper investigation. However, as the Ten-Item Inventory is only a "snapshot" measure of individual differences in personality, any conclusions made at this stage would be extremely limited. As such, we highlight the possibility for further research into personality theory and consistent/inconsistent self avatar representations affecting a user's perception of others.

Overall results suggested that there were significant values for Subjective Trust amongst participants in the consistent condition AVxAV and JCxJC over inconsistent conditions AVxJC and JCxAV, supporting H1 (see **Figure 10**). This result might be explained by several factors. For example, perhaps having the

FIGURE 9 | Scatter plot of BIG FIVE personalities vs. social presence factors. *$p < 0.05$.
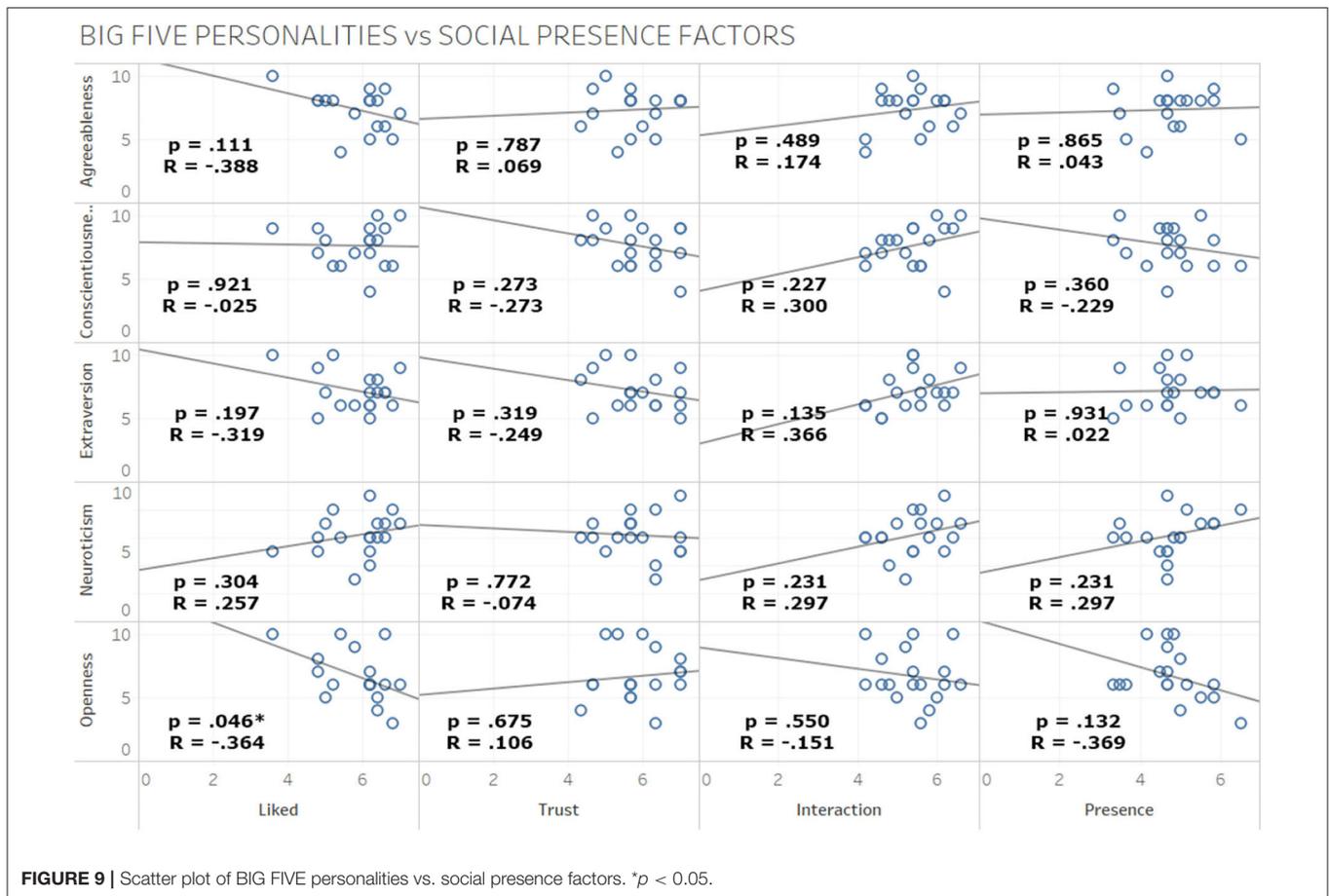
same self-representation fostered higher levels of *Social Presence*, leading to increased interpersonal trust between participants. Or alternatively, the consistency made finding ways to express themselves non-verbally easier and less of a cognitive effort.

Surprisingly, mean Liked scores were observed to be higher in inconsistent conditions than consistent conditions. There are some studies that have come to show positive correlation patterns between *Liked* and *Subjective Trust*, but there are also those that do not. A study found that being mimicked did not change trust or liking within or across CVE social groups (Hale, 2017).
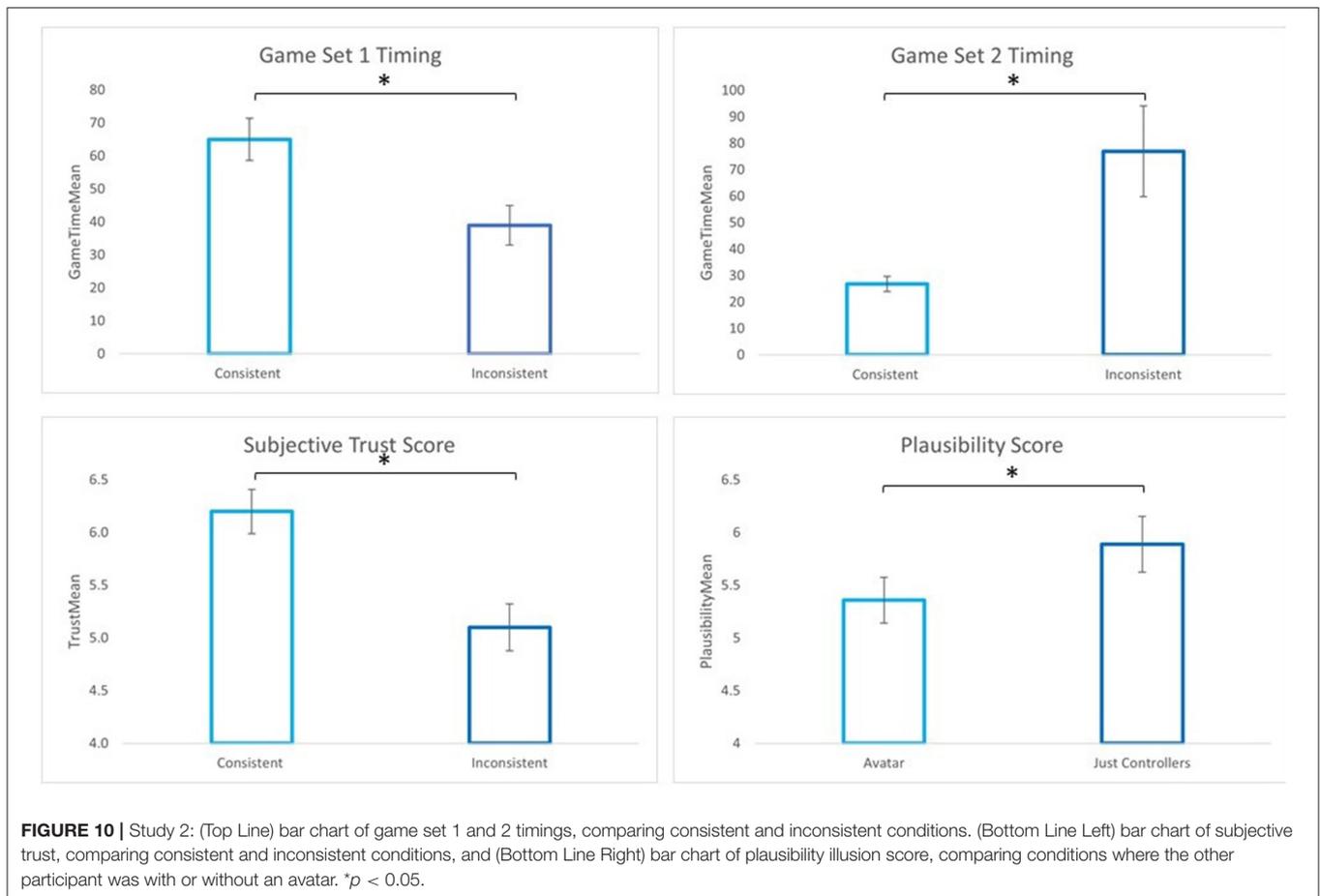
Results for *Plausibility* was higher in conditions where the "other" participant did not have an avatar (see **Figure 10**). Potentially, this could be due to technical limitations when engaging with the environment, e.g., that the avatar rendered was not realistic enough and therefore has hindered the *Plausibility Illusion* rather than facilitating it.

5. GENERAL DISCUSSION

This work extends the research introduced in previous work (Pan and Steed, 2017) by focusing on the impact of self-representation and consistency in CVE. As we continue to progress within this virtual age it is important to understand the effect of consistency in avatar representation to inform the development of social collaborative applications within the various industries utilizing

VR. The results of this investigation firstly reinforces the positive effect of the Self Avatar within social interactions but moreover suggests consistency can improve trust if there is an equal and transparent dynamic between active participants. This is highlighted in Study 2, where we see subjective scores are higher in consistent conditions. This is also true for productivity. Study 1 highlights a potential caveat in utilising a confederate in paired studies which is supported by previous literature (Martin, 1970; Feng et al., 2004; Kuhlen and Brennan, 2013). When using a confederate who is acting in deceit, it invites suspicion into the social dynamic which may affect interactions between pairs. In this study it is suggested in particular that when a confederate is deceitful and using a self avatar, this may have a negative effect on subjective levels of trust. This may be due to greater non-verbal "leakage" of social signals through the avatar, that enable the participant to pick up more cues of deceit. This shows the potential difficulties of using experimental designs based on confederates.

Using a social dilemma exercise to gather objective measures of trust proved to be unreliable in this context. We see completely opposing results in both *Subjective Trust* and *Liked* between Study 1 and Study 2. This also could potentially have been affected, or compounded by the use of a confederate. In Study 1, the confederate is an "expert" at the experiment process and therefore has less cognitive load overcoming the learning curve

FIGURE 10 | Study 2: (Top Line) bar chart of game set 1 and 2 timings, comparing consistent and inconsistent conditions. (Bottom Line Left) bar chart of subjective trust, comparing consistent and inconsistent conditions, and (Bottom Line Right) bar chart of plausibility illusion score, comparing conditions where the other participant was with or without an avatar. *$p < 0.05$.

of using the system and working in a pair to complete the task. In Study 2, both are novice participants to the system and perhaps in this case it was more difficult to establish relationships whilst trying to complete the task correctly. Alternatively, in this context participants may have found their partners trustworthy to complete the task but not likable. The type of trust and likability that would warrant sharing something as valuable as money perhaps had not been able to develop. In Study 1, the confederate played a "consistent role" which may have helped participants to relate to them better.

Study 2 showed that the efficiency of consistent and inconsistent pairs varied over time. Initially, inconsistent pairs were faster, possibly due to one partner naturally taking on a leadership role. However, over time the consistent pairs were more efficient, perhaps because they were able to establish more effective collaboration strategies after an initial period of familiarisation with each other.

In this study we looked at the effect of having consistent and inconsistent conditions between partners when using a confederate and when using paired participants, as this could have interesting implications in the design of shared virtual spaces—and our findings have both supported and challenged previous notions. More importantly, this approach has given insight into how we can begin thinking about consistency in

utilizing the Self Avatar. More research needs to be done in this area to get a fuller understanding of this phenomenon.

6. FUTURE WORK

Future work will consist of working with a larger sample size. It would also be interesting to gather gaze and arm movement data from the participants whilst they are playing in the different conditions. We also wish to further investigate the effect of avatar appearance, for example, adding more diversity in skin tone and playing with consistency in first person perception.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by (Study1) George Mason Ethics Committee (USA)/(Study2) Prof. Robert Zimmer, Department of Computing, Goldsmiths University of London (UK). The

participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

TC-W was the primary author and made a substantial contribution in the conception and design of the work (involved in initial design and implementation, and was in charge of recruitment and ethics submission) and also analysing and interpreting the data (produced all graphs and ran most of the analysis for the data). TC-W and ZO'S ran the experiment with participants and ZO'S also made a substantial contribution to data acquisition and the manuscript (sections of general edits and proofreading). XP made a substantial contribution to the conception and design of the work (involved in the initial design, helped with the ethics application), data analysis (helped with the data analysis and the plots), as well as drafting the manuscript (wrote part of sections of Introduction and Results). MG made a substantial contribution to drafting the manuscript (sections of the Discussion) and contribution to data interpretation. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frvir.2021.648601/full#supplementary-material

REFERENCES

Bailenson, J. N., Swinth, K., Hoyt, C., Persky, S., Dimov, A., and Blascovich, J. (2005). The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence Virt. Augment. Reality* 14, 379–393. doi: 10.1162/105474605774785235

Botvinick, M., and Cohen, J. (1992). Rubber hands 'feel' touch that eyes see. *Nature* 391:756. doi: 10.1038/35784

Dey, A., Piumsomboon, T., Lee, Y., and Billinghurst, M. (2017). "Effects of sharing physiological states of players in collaborative virtual reality gameplay," in *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, CO), 4045–4056. doi: 10.1145/3025453.3026028

Doyen, S., Klein, O., Pichon, C., and Cleeremans, A. (2012). Behavioral priming: it's all in the mind, but whose mind? *PLoS ONE* 7:e29081. doi: 10.1371/journal.pone.0029081

Feng, J., Lazar, J., and Preece, J. (2004). Empathy and online interpersonal trust: a fragile relationship. *Behav. Inform. Technol.* 23, 97–106. doi: 10.1080/01449290310001659240

Garau, M., Slater, M., Bee, S., and Sasse, M. A. (2001). "The impact of eye gaze on communication using humanoid avatars," in *CHI '01: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, WA), 309–316. doi: 10.1145/365024.365121

George, C., Eiband, M., Hufnagel, M., and Hussmann, H. (2018). "Trusting strangers in immersive virtual reality," in *IUI '18 Companion: Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion* (Tokyo), 1–2. doi: 10.1145/3180308.3180355

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., and Soutter, C. L. (2000). Measuring trust. *Q. J. Econ.* 115, 811–846. doi: 10.1162/003355300554926

Gong, L., and Nass, C. (2007). When a talking-face computer agent is half-human and half-humanoid: human identity and consistency preference. *Hum. Commun. Res.* 33, 163–193. doi: 10.1111/j.1468-2958.2007.00295.x

Hale, J. (2017). *Using novel methods to examine the role of mimicry in trust and rapport* (Ph.D. thesis), University College London, University of London, London, United Kingdom.

Ijsselsteijn, W., de Kort, Y., and Poels, K. (2013). *The Game Experience Questionnaire*. Eindhoven: Technische Universiteit Eindhoven.

Johnson, N. D., and Mislin, A. A. (2007). "Multiview: improving trust in group video conferencing through spatial faithfulness," in *CHI '07: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, CA), 1465–1474. doi: 10.1145/1240624.1240846

Johnson, N. D., and Mislin, A. A. (2011). Trust games: a meta-analysis. *Presence Virt. Augment. Reality* 32, 865–889. doi: 10.1016/j.joep.2011.05.007

Kilteni, K., Bergstrom, I., and Slater, M. (2013). Drumming in immersive virtual reality: the body shapes the way we play. *IEEE Trans. Vis. Comput. Graph.* 19, 597–605. doi: 10.1109/TVCG.2013.29

Kilteni, K., Groten, R., and Slater, M. (2012). The sense of embodiment in virtual reality. *Presence Teleoper. Virt. Environ.* 21, 373–387. doi: 10.1162/PRES_a_00124

Kim, K., Bolton, J., Girouard, A., Cooperstock, J., and Vertegaal, R. (2012). "Telehuman: effects of 3D perspective on gaze and pose estimation with a life-size cylindrical telepresence pod," in *CHI 12: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, TX), 2531–2540. doi: 10.1145/2207676.2208640

Kondo, R., Sugimoto, M., Minamizawa, K., Hoshi, T., Inami, M., and Kitazaki, M. (2018). Illusory body ownership of an invisible body interpolated between virtual hands and feet via visual-motor synchronicity. *Sci. Rep.* 8:7541. doi: 10.1038/s41598-018-25951-2

Kuhlen, A. K., and Brennan, S. E. (2013). Language in dialogue: when confederates might be hazardous to your data. *Psychonom. Bull. Rev.* 20, 54–72. doi: 10.3758/s13423-012-0341-8

Latoschik, M. E., Roth, D., Gall, D., Achenbach, J., Waltemate, T., and Botsch, M. (2017). "The effect of avatar realism in immersive social virtual realities," in *VRST '17: Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology* (Gothenburg), 1–10. doi: 10.1145/3139131.3139156

Martin, D. J. (1970). Suspicion and the experimental confederate: a study of role and credibility. *Sociometry* 33, 178–192. doi: 10.2307/2786328

Masahiro, M. (1970). Bukimi no tani [The uncanny valley]. *Energy* 7, 33–35.

Maxwell, S. E., Delaney, H. D., and Kelley, K. (2004). *Designing Experiments and Analyzing Data: A Model Comparison Perspective, 2nd Edn*. New Jersey, NJ: Lawrence Erlbaum Associates.

Mohler, B. J., Creem-Regehr, S. H., Thompson, W. B., and Büthoff, H. H. (2010). The effect of viewing a self-avatar on distance judgments in an

hmd-based virtual environment. *Presence Virt. Augment. Reality* 19, 230–242. doi: 10.1162/pres.19.3.230

Murphy, D. (2017). "Bodiless embodiment: a descriptive survey of avatar bodily coherence in first-wave consumer VR applications," in *2017 IEEE Virtual Reality (VR)* (Los Angeles, CA). doi: 10.1109/VR.2017.7892278

Myerson, R. B. (1992). On the value of game theory in social science. *Rational. Soc.* 4, 62–73. doi: 10.1177/1043463192004001008

Oh, C. S., Bailenson, J. N., and Welch, G. F. (2018). A systematic review of social presence: definition, antecedents, and implications. *Front. Robot. AI* 5, 413–434. doi: 10.3389/frobt.2018.00114

Pan, X., Gillies, M., and Slater, M. (2015). Virtual character personality influences participant attitudes and behavior—an interview with a virtual human character about her social anxiety. *Rational. Soc.* 2:1. doi: 10.3389/frobt.2015.00001

Pan, Y., and Steed, A. (2017). The impact of self-avatars on trust and collaboration in shared virtual environments. *PLoS ONE* 12:e0189078. doi: 10.1371/journal.pone.0189078

Rae, I., Takayama, L., and Mutlu, B. (2013). "In-body experiences: embodiment, control, and trust in robot-mediated communication," in *CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris), 1921–1930. doi: 10.1145/2470654.2466253

Rammstedt, B., and John, P. O. (2007). Measuring personality in one minute or less: a 10-item short version of the big five inventory in English and German. *J. Res. Pers.* 41, 203–212. doi: 10.1016/j.jrp.2006.02.001

Regenbrecht, H., Haller, M., Hauber, J., and Billinghurst, M. (2006). Carpeno: Interfacing remote collaborative virtual environments with table-top interaction. *Virt. Reality* 10, 95–107. doi: 10.1007/s10055-006-0045-3

Skarbez, R., Neyret, S., Brooks, F. P., Slater, M., and Whitton, M. C. (2017). A psychophysical experiment regarding components of the plausibility illusion. *IEEE Trans. Vis. Comput. Graph.* 23, 1369–1378. doi: 10.1109/TVCG.2017.2657158

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *R. Soc.* 364, 3549–3557. doi: 10.1098/rstb.2009.0138

Slater, M., and Steed, A. (2000). A virtual presence counter. *Presence* 9, 413–434. doi: 10.1162/105474600566925

Steed, A., Pan, Y., Zisch, F., and Steptoe, W. (2016). "The impact of a self-avatar on cognitive load in immersive virtual reality," in *2016 IEEE Virtual Reality (VR)* (Greenville, SC). doi: 10.1109/VR.2016.7504689

Steed, A., Slater, M., and Sadagic, A. (1999). "Leadership and collaboration in shared virtual environments," in *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)* (Houston, TX). doi: 10.1109/VR.1999.756941

Verberne, F., Ham, J., Ponnada, A., and Midden, C. (2013). "Trusting digital chameleons: the effect of mimicry by a virtual social agent on user trust," in *8th International Conference on Persuasive Technology (PERSUASIVE 2013)* (Berlin; Heidelberg), 234–245. doi: 10.1007/978-3-642-37157-8_28

Witmer, B. G., and Singer, M. J. (1998). Measuring presence in virtual environments: a presence questionnaire. *Presence Virt. Augment. Reality* 7, 225–240. doi: 10.1162/105474698565686

Yee, N., and Bailenson, J. (2007). The proteus effect: the effect of transformed self-representation on behavior. *Hum. Commun. Res.* 33, 271–290. doi: 10.1111/j.1468-2958.2007.00299.x

Yoon, B., Kim, H., Lee, G. A., Billinghurst, M., and Woo, W. (2019). "The effect of avatar appearance on social presence in an augmented reality remote collaboration," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (Osaka), 547–556. doi: 10.1109/VR.2019.8797719