Realistic Motion Avatars are the Future for Social Interaction in Virtual Reality

Shane L. Rogers *, Rebecca Broadbent, Jemma Brown, Alan Fraser and Craig P. Speelman

Cognition Research Group, Edith Cowan University, Perth, WA, Australia

This study evaluated participant self-reported appraisal of social interactions with another person in virtual reality (VR) where their conversational partner was represented by a realistic motion avatar. We use the term realistic motion avatar because: 1. The avatar was modelled to look like the conversational partner it represented, and 2. Full face and body motion capture was utilised so that the avatar mimicked the facial and body language of the conversational partner in real-time. We compared social interaction in VR with face-to-face interaction across two communicative contexts: 1. Getting acquainted conversation, and 2. A structured interview where the participant engaged in self-disclosure about positive and negative experiences. Overall, participants largely indicated they preferred face-to-face over VR communication. However, some participants did indicate a preference for VR communication. Additionally, an analysis of post-conversation ratings indicated no significant difference for rated enjoyment, understanding, self-disclosure, comfort, and awkwardness between communication modes. The only ratings where face-to-face was found to be superior was for perceived closeness across both types of communication, and for feeling understood specifically when disclosing negative experiences. Most participants perceived frequent eye contact in both face-to-face and VR interaction, but typically more eye contact when face-to-face. Eye contact was positively associated with rated enjoyment, closeness, and comfort. Overall, our findings suggest that harnessing full face and body motion capture can make social interaction in VR very similar to face-to-face interaction. We anticipate that VR social interaction is poised to become the next major technological evolution for human computer mediated communication and suggest avenues for further research.

Keywords: virtual reality, avatar, motion capture (Mocap), face-to-face (dyadic) communication, computer mediated communication (CMC), eye contact, self-disclosure, social interaction

1 INTRODUCTION

Computer-mediated communication has generally increased during the recent COVID-19 pandemic (Meier et al., 2021; Pfund et al., 2021; Rogers & Cruickshank, 2021). Virtual reality (VR) is next on the horizon to become a popular computer-mediated mode of social interaction (Gunkel et al., 2018; Seymour et al., 2021). In this paper we define VR specifically as an immersive three-dimensional experience in a virtual world *via* a head-mounted display (HMD). We investigated participant experience of social interaction in virtual reality where the face and body of the conversational partner's avatar is controlled in real-time *via* motion capture. We contrast participant experience of social interaction in VR with face-to-face interaction across two social contexts—getting acquainted

conversation, and in a semi-structured interview where the participant discloses positive and negative personal experiences. The primary aim of this research is to investigate if full face and body motion capture enables social interaction in VR to more closely approximate face-to-face interaction than what has previously been possible.

1.1 Basic Avatars Versus Realistic Motion Avatars in Virtual Reality

When conceptualizing social interaction in virtual reality Pan and Hamilton (2018) have pointed out that it is useful to consider two important dimensions—Interaction dynamics (i.e., extent and responsiveness of verbal and non-verbal feedback), and graphical realism. Historically a major limitation of VR as a form of real-time computer-mediated communication has been that the digital representation of self and others consists of non-expressive avatars that lack realistic body and facial expression. In other words, such avatars have a low level of interaction dynamics with low graphical realism. We refer to this type of avatar as a *basic avatar*.

The advancement of VR technology and computing hardware has resulted in the release of several VR platforms for online social interaction over recent years. Some examples of the current largest players in this space are *VRChat, Altspace, Rec Room, and Facebook Horizon*. In these platforms the avatar's body reacts based on the user's movements *via* motion tracking from the headset and controllers. The avatar mouth also moves to simulate talking when the player talks. For those who own head mounted displays (HMD's) with in-built eye tracking some platforms harness the eye tracking to produce more responsive eye movements of the avatar. However, the fluidity of body and face movements remains limited by rudimentary motion tracking within the current hardware. We therefore refer to the current state of avatars in these VR social platforms as *semi-realistic motion avatars*. We argue that the semi-realistic nature of the avatars has been holding back wider adoption. For example, at the time of writing this article according to *steamcharts. com* VRChat average concurrent users gradually rose across 2020 from 7,973 (Jan 2020) to 14,910 (Jan 2021). However usage remained relatively stable during the first half of 2021 with 14,328 concurrent players in June 2021.

To move from semi-realistic motion avatars to what we will call *realistic motion avatars* we argue requires more sophisticated motion tracking than is the current norm in social VR applications. Social information communicated *via* both facial expression and body posture can be based on subtle movements (Karaaslan et al., 2020; Maloney et al., 2020; Meeren et al., 2005; Vesper & Sevdalis, 2020). For example, a slight smile combined with slight eyebrow raise can help to communicate when a message is intended to be interpreted as sarcasm (Attardo et al., 2003). Therefore, for social interaction in VR to feel more immersive and to be experienced as more like natural face-to-face interaction avatars need to mimic the human actor in a subtle and naturalistic fashion.

A recent study by Seymour et al. (2021) examined perceptions of audience members who acted as third-party observers of social interactions of motion-captured realistic avatars in HMD VR compared with viewing on a screen. The observed social interactions consisted of interviews between the host (acted by M. Seymour) and guests consisting of experts from the visual design and/or movie industries. The avatar of the host was a photo-realistic graphical recreation animated in real-time *via* full facial motion capture. The face of the guest avatars was custom designed to resemble the person but were not at the same photo-realistic level as the host. The guest avatars were animated in real-time *via* approximate facial animation using deep-learning extrapolation from capture of the mouth and eye positions alone. The participants (both guests, and observers) gave questionnaire and qualitative feedback that indicated they felt very positive about the interactions and were more positive about the interaction experienced *via* HMD VR compared to watching *via* a computer monitor. Participants also consistently indicated a preference for the photo-realistic avatar of the host over the guest avatars of lower graphical fidelity. The study by Seymour et al. (2021) provides some good preliminary evidence of positive perceptions of realistic motion avatars.

1.2 The Present Study

Researchers have begun to analyze participant experiences interacting in HMD-based virtual reality with motion-captured avatars in real-time (Latoschik et al., 2017; Pettersson & Sundstedt, 2017; Smith & Neff, 2018; Wu et al., 2019; Seymour et al., 2021; Wu et al., 2021). However, to date no studies have examined experience interacting with a full face *and* body motion captured avatar. In the present study we aim to explore participant experience when interacting with a realistic motion avatar (i.e., animated *via* full face and body motion tracking in real-time) in VR compared to analogous interactions face-to-face. We explore two broad communicative contexts—getting acquainted conversation, and a semi-structured interview where the participant is required to disclose positive and negative experiences. Our interest in the former is due to the increased role VR will likely play in the future for online casual social interaction (Gunkel et al., 2018; Seymour et al., 2018; Smith & Neff, 2018; Maloney et al., 2020; Seymour et al., 2021), and in the latter due to the potential of VR as an alternative method of communication for psychological therapy (Rehm et al., 2016; Baccon et al., 2019; Slater et al., 2019; Pedram et al., 2020; Geraets et al., 2021; Pimentel et al., 2021; Sampaio et al., 2021).

In the present study after each interaction participants were asked to rate their experience regarding enjoyment, closeness, self-disclosure, comfort, and awkwardness. This methodological approach and specific items were inspired by prior research conducted by Sprecher and others on people rating their experience of brief social interactions (Sprecher, 2014; Sprecher, 2021; Sprecher & Hampton, 2017; Sprecher et al., 2016). Like this prior research, we were anticipating that participants would overall report a high level of enjoyment, closeness, feeling understood, self-disclosure, and comfort with their partner during face-to-face interactions (and low awkwardness). We were anticipating that face-to-face interaction would be rated more positively than VR

TABLE 1 | Participant self-reported social engagement across different modes of interaction, and experience playing computer games over the prior year.

	None	A few times	About once a month	About once a week	Multiple times per week
Face-to-face conversations	0 (0%)	0 (0%)	0 (0%)	0 (0%)	52 (100%)
Phone conversations	0 (0%)	1 (2%)	0 (0%)	9 (17%)	42 (81%)
Screen-based conversations	0 (0%)	19 (37%)	9 (17%)	13 (25%)	11 (21%)
Conversations in virtual reality	51 (98%)	1 (2%)	0 (0%)	0 (0%)	0 (0%)
Screen-based computer games	16 (31%)	14 (27%)	6 (12%)	5 (10%)	11 (21%)
Virtual reality computer games	42 (81%)	7 (14%)	3 (6%)	0 (0%)	0 (0%)

interaction and be the preferred mode of interaction for most participants. We expected this because face-to-face allows for a richer non-verbal experience and because it is the most-engaged in and familiar experience (Baltes et al., 2002; Kock, 2005; Vlahovic et al., 2012; Sprecher, 2014; Sprecher & Hampton, 2017). However, due to the use of full body and face motion capture for the VR avatar we were anticipating that the difference in ratings between the face-to-face and VR modes of interaction might only be small.

Some people might prefer technology-mediated communication (such as text-chat, video chat, or interaction in virtual reality) over face-to-face interaction for therapy (Suler, 2000; Falconer et al., 2019). One reason underlying this is that some individuals can feel more comfortable disclosing about sensitive topics when there is a greater sense of interpersonal distance between self and therapist (Suler, 2000). Therefore, in the present study we hypothesize that some of our participants will report a preference for the VR interaction over face-to-face interaction, especially when disclosing negative experiences compared to disclosure of positive experiences.

Finally, we also asked participants to estimate the amount of eye contact that occurred after each interaction. Eye contact is an important aspect of face-to-face interaction as the eyes contribute to the communication of emotion (Baron-Cohen et al., 1997). Side to side eye movements on and off face can help to regulate turn-taking which helps to make face-to-face conversation feel more fluid and comfortable (Ho et al., 2015; Rogers et al., 2018). In neuro-typical populations in Western cultures the extent of perceived eye contact has been found to be positively associated with conversational enjoyment (Kleinke, 1986; Akechi et al., 2013; Senju & Johnson, 2009). Therefore, in our study we felt it was important to include an exploration of the perception of eye contact during the VR social interaction compared with face-to-face interaction.

As per previous research we hypothesized that for face-to-face interactions eye contact would be perceived as occurring for a high percentage of the conversation (Rogers et al., 2019; Rogers et al., 2018), and greater perceived eye contact would be associated with greater conversational enjoyment (Akechi et al., 2013; Senju & Johnson, 2009). Our comparison between the extent of perceived eye contact across the face-to-face and VR modes of interaction was largely exploratory. People exhibit a bias to perceive gaze that is directed somewhere towards the face region as eye contact (Rogers et al., 2019). Therefore, we were expecting that the perception of eye contact in the VR interaction would be relatively high. However, the precise extent that participants would report experiencing eye contact in VR social interactions was largely unknown.

2 METHODS

2.1 Participants
Fifty-two under-graduate psychology students participated in the study (10 Males and 42 females, Mean age = 31.73 years, SD = 10.57, Min = 18, Max = 53). As can be seen in **Table 1**, these participants were generally socially active people, with all participants engaging in face-to-face conversations multiple times per week, and 81% engaging in phone conversations multiple times per week. While most had some experience with computer games over the past year (69%), a minority had prior experience with virtual reality (20%). Ethics approval for this research was granted by the Edith Cowan University ethics committee (REF: 2019–01013-ROGERS).

2.2 Materials: Hardware
To run the VR experience a desktop computer was used with the following specifications: CPU (Intel i7-9700K), GPU (Nvidia RTX 2080), RAM (32GB DDR4), Storage (500GB SSD and 2TB HDD). The VR head mounted display (HMD) unit used was an Oculus Rift S. This is a tethered headset with 1280x1440 resolution per eye with an 80Hz refresh rate that connects to the PC *via* DisplayPort. For body motion capture the Perception Neuron Studio System was used. This comprises of 17 inertial sensors for the body, and an additional 12 sensors with the Perception Neuron Studio Gloves. The sensors communicate (<20 millisecond latency) with a dedicated transceiver that can connect to the PC *via* USB or ethernet. For facial tracking we used the iclone LIVE FACE app that requires an iPhone 10 or above (we used an iPhone 11) that we mounted in front of the actor's face using a FACEWARE indie headcam helmet. The iPhone app had the potential to communicate wirelessly to the PC but we ran it *via* a USB connection to ensure the best possible latency.

2.3 Materials: Software
The VR simulation was run *via* Unreal Engine 4 (UE4). A basic virtual room was created that all VR interactions took place in, see **Figure 1A**. During the interactions, the participant was positioned in the room on the side of the desk facing the VR avatar, see **Figure 1B**. The participant did not possess an avatar of their own.
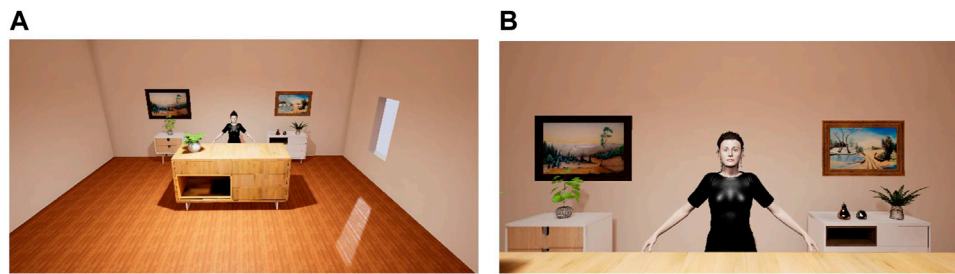
FIGURE 1 | The virtual space used in the present study from **(A)** a bird's eye perspective, and **(B)** the perspective of the participant.
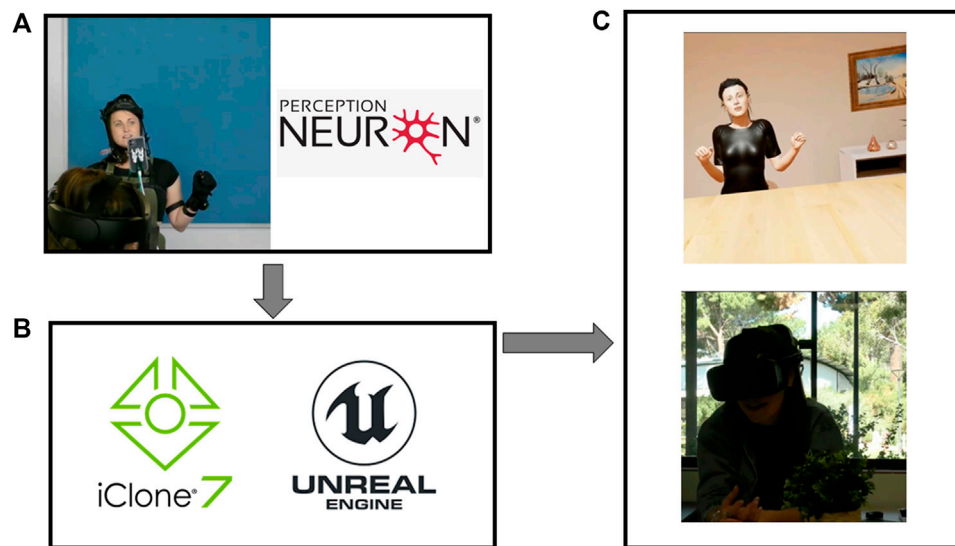

FIGURE 2 | The methodological pipeline used in the present study. **(A)** The conversational partner of the participant was physically in front of the participant having their body movement tracked by a Perception Neuron motion capture system, and their face tracked *via* the LIVE FACE app *via* an iPhone mounted in front of their face. **(B)** The motion capture data was fed to the software iclone 7 and UNREAL engine to enable the avatar of the conversational partner to mimic face and body movements in real-time. **(C)** This avatar was broadcast into the virtual space in the UNREAL engine so that the participant could interact with them *via* an Oculus head mounted display. Note that the avatar **(C)** was modelled to physically resemble the real-life conversational partner **(A)**. For an example of what the interaction looked like in practice see: https://youtu.be/rrdQ3Qio2WQ.

Motion capture of the body and hands was achieved *via* data from the Perception Neuron inertial motion capture sensors wirelessly sent to the dedicated receiver that sends the data to the Perception Neuron Axis Studio software. This data is then mapped onto the avatar in iclone 7 software *via* the Motion Live plug-in. The motion capture of the face was achieved *via* the LIVE FACE app on the iPhone 11 mounted in front of the actor's face that sent data directly to the iclone software *via* USB cable (*via* Motion Live plugin in iclone). The avatar is directly linked in real-time between the iclone software and UNREAL software *via* the Live Link plug-in for UNREAL. By having the participant wear the Oculus HMD and using the preview function in UNREAL, the participant could see the character in virtual reality with body and face movements mimicking the actor wearing the motion capture equipment. For a graphical depiction of this pipeline see **Figure 2**. An additional piece of software used was iclone

Character Creator with the Headshot plug-in to make the avatars resemble the people who the avatars represented in this study. Participants met the people in control of the avatars not only in VR, but also face-to-face in the lab during this study. If the avatars did not resemble their controller this likely would have been a jarring experience for participants which may have artificially diminished the enjoyment of the VR experience. Therefore, in the present study we felt it was important that the avatars looked very similar to the people they were created to represent.

2.4 Procedure

The study procedure contained two phases, a getting acquainted phase, and a self-disclosure phase. These are described separately below. Prior to having the conversations participants filled out a brief preliminary survey where they indicated their age, gender,

and the extent that over the past year they had conversed with others *via* different communication modes: Face-to-face, phone, screen-based, and VR. Additionally, participants were asked to recall two positive and two negative life experiences that they would be asked about later. They were required to write a very brief description of each experience and provide a rating of how bad (for negative experiences) or good (for positive experiences) they recalled feeling during the time the experience occurred. This was rated on a 4-point scale of 1) Slightly 2) Somewhat 3) Very 4) Extremely. Overall, the proportion of participant responses for "bad" ratings for negative disclosures were: Slightly (6%), Somewhat (23%), Very (36%), and Extremely (36%). The proportion of participant responses for "good" ratings for positive disclosures were: Slightly (0%), Somewhat (3%), Very (29%), and Extremely (68%). No significant difference was found between face-to-face or VR disclosures for either negative ($z = 0.99$, $p = 0.32$) or positive ($z = 0.47$, $p = 0.64$) experiences. Therefore, in the present study the intensity of the negative and positive experiences participants chose to disclose was considered equal across face-to-face and VR modes. This indicates that intensity of positive/negative experiences did not act as a confounding variable.

2.4.1 Getting Acquainted Phase

Participants started off the study by having two 4-min getting acquainted conversations with two of the study investigators (RB and JB). Both RB and JB were females in their 20 s at the time of the study. Participants were randomly allocated to converse initially in either face-to-face or VR modes of communication. The initial conversational partner (RB & JB) was also randomly allocated across participants. Both RB and JB endeavored to behave similar to each other and be as consistent as possible across different interactions. For example, by wearing the same set of clothes for all interactions, trying to keep their body language and tone of voice consistent, and asking some consistent initial questions such as "so, what do you study?" and "how are you enjoying your studies so far?". Immediately after each interaction the participant filled out a brief survey of their experience. They were asked to rate the conversation regarding extent of enjoyment, a sense of closeness to their partner, feeling understood, their self-perceived amount of self-disclosure, comfort, and awkwardness. These items were rated on a 4-point scale: 1) Not at all 2) A little bit 3) Quite a bit 4) A lot. Participants were also asked to rate their self-perception of the extent of eye contact between themself and their partner on a 10-point scale that ranged from 10 to 100% increasing in 10% increments.

After the second interaction participants were asked to reflect on both interactions and indicate which mode of communication they most enjoyed, felt closest to their partner, disclosed the most information about themselves, felt the most comfortable, felt the most awkward, and in which mode they found it easiest to relax and be themselves. They indicated this *via* forced-choice options: face-to-face, VR, or about the same.

During the face-to-face chat, participants sat directly across from the experimenter they were conversing with at a small table. During the VR chat, the avatar of the experimenter was in a

FIGURE 3 | A photo showing the lab set up which provides an indication of the physical distance between the participant (left) with their conversation partner (on the right). For face-to-face conversations the conversational partner was seated across from the participant at the table. The chair that the conversational partner sat in can be seen at the other side of the table from the participant in this photograph. During VR interactions the participant interacted with the avatar which form the participant perspective appeared similarly to **Figure 1B**.

standing position slightly back from the table, rather than sitting in the exact type of position as the face-to-face chat. This was done because after some pre-testing it became clear that this was the easiest/fastest way to set things up and avoided any clipping issues. Clipping is a term used to describe how computer avatars might have body parts pass through other parts of their body or inanimate objects, such as a table. In the future we plan on assessing different positional configurations for the avatar, but in this initial study we decided to use the simplest approach. The participant perspective of the VR avatar can be seen in **Figures 1**, **2**. In **Figure 3** we show how the set-up looked from outside of VR.

2.4.2 Self-Disclosure Phase

After rating the initial getting acquainted conversations participants were moved on to the next phase that involved a semi-structured interview with one of the experimenters (RB). In both face-to-face and VR modes, the participant was asked by RB about one negative and one positive experience. RB endeavored to keep her behavior as consistent as possible across the participants regarding her body language and tone of voice. A decision was made not to counterbalance the order of disclosures so that the study could finish with the participants disclosing a positive event and thus finish their participation in the study on a positive note. Therefore, the order of disclosures was always negative—positive—negative—positive. For each disclosure participants were asked the same set of questions in order: 1. What the event was, 2. When it happened, 3. How it happened, and 4. How it made you feel.

The order of the mode of interaction that was engaged in first was randomised across participants (i.e., face-to-face or VR). The rating of the experience of conversing about each disclosure was conducted at the end of each pair of disclosures (i.e., a negative/positive pair). For example, if the participant engaged in the face-to-face exchange initially, they would engage in the structured interview about the negative experience, then the positive
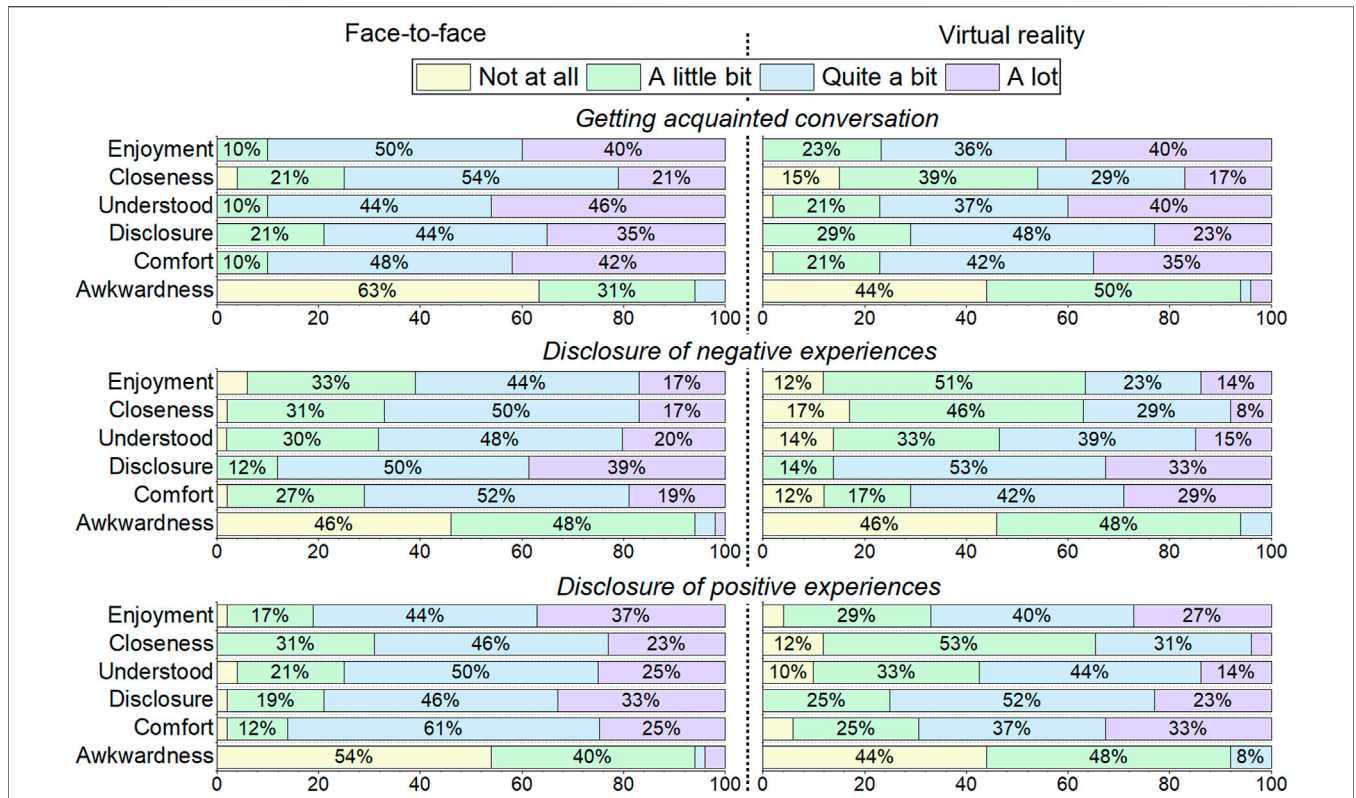
FIGURE 4 | Participant ratings of their experience conversing across different communicative contexts both face-to-face, and in virtual reality. Note that participants made the ratings directly after each of the experiences.

experience, then make their ratings. Subsequently they would engage in a semi-structured interview with the VR avatar about the other negative and positive experiences, then rate those experiences.

The ratings of the experiences were the same ratings used in the getting acquainted phase. Participants provided ratings for each one of their disclosures (i.e., 4 in total). At the very end of the study, participants were asked to directly contrast which mode they liked the best across the same dimensions asked in the getting acquainted phase. One additional item asked participants to indicate which communication mode they would most prefer if they were to talk to a therapist.

3 RESULTS

The raw data for this article is available on Figshare: https://doi. org/10.6084/m9.figshare.15085749.v1.

3.1 Rating Conversational Experience

After face-to-face and virtual reality conversations participants rated their experience on several adjectives, see **Figure 4**. Due to the ordinal nature of the response scale we conducted a series of Wilcoxon signed rank tests when comparing across face-to-face and VR modes, with a Bonferroni-adjustment to the accepted p-value (i.e., 6 comparisons: 0.05/6 = 0.008) (Field, 2013).

For getting acquainted conversations there was no significant difference found between face-to-face and VR conversations for ratings of enjoyment, feeling understood, self-disclosure, comfort, and awkwardness (all $zs \leq 2.13$, $ps > 0.008$). In both face-to-face and VR modes across the adjectives of enjoyment, feeling understood, disclosure and comfort ratings of *quite a bit - a lot* ranged from 71 to 90%. For awkwardness ratings, 95 and 94% rated the conversation *not at all - a little* awkward in both face-to-face and VR modes. The only adjective where a significant difference was found was for closeness ($z = 3.18$, $p < 0.001$), with 75% rating *quite a bit - a lot* when face-to-face versus 46% in VR.

For the disclosure of negative experiences, no significant differences were found between face-to-face and VR across the adjectives of disclosure and comfort ratings (all $zs \leq 0.73$, $ps > 0.008$). For both face-to-face and VR ratings of *quite a bit - a lot* ranged from 71 to 89%. For awkwardness ratings, 94% rated the conversation *not at all - a little* awkward in both face-to-face and VR, with no significant difference. Significant differences were found for the adjectives of feeling understood ($z = 3.40$, $p = 0.001$), and closeness ($z = 4.20$, $p < 0.001$). For feeling understood, 89% rated *quite a bit - a lot* when face-to-face compared to 54% in VR. For closeness, 67% rated *quite a bit - a lot* when face-to-face compared to 37% in VR. Face-to-face was rated more enjoyable as 61% rated *quite a bit - a lot* compared with 37% in VR, however this just failed to
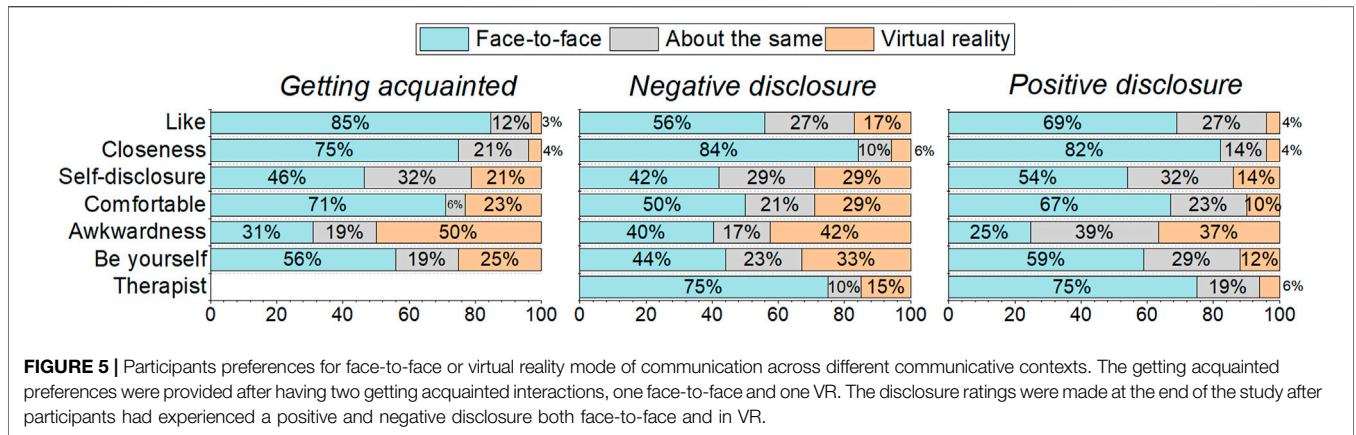
FIGURE 5 | Participants preferences for face-to-face or virtual reality mode of communication across different communicative contexts. The getting acquainted preferences were provided after having two getting acquainted interactions, one face-to-face and one VR. The disclosure ratings were made at the end of the study after participants had experienced a positive and negative disclosure both face-to-face and in VR.
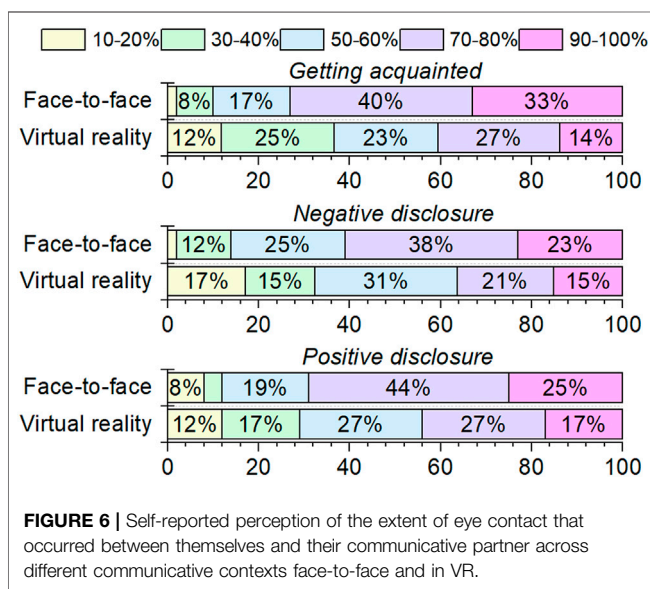

FIGURE 6 | Self-reported perception of the extent of eye contact that occurred between themselves and their communicative partner across different communicative contexts face-to-face and in VR.

reach the 0.008 threshold for statistical significance ($z = 2.46$, $p = 0.01$).

For the disclosure of positive experiences differences between face-to-face and VR mirrored the results for the disclosure of negative experiences. Participants typically reported feeling a greater sense of closeness to their partner ($z = 4.53$, $p < 0.001$) and more understood ($z = 2.93$, $p = 0.003$) when disclosing face-to-face compared with VR. There was no significant difference for disclosure, comfort, and awkwardness (all $zs \leq 1.29$, $ps > 0.008$). The ratings for enjoyment approached the significance threshold ($z = 2.18$, $p = 0.03$).

While the post-conversational ratings suggest only minimal differences in experience between face-to-face and VR interaction, when asked to indicate a preference the results more clearly favour face-to-face interaction, see **Figure 5**. Across each type of interaction most participants liked face-to-face more, felt closer to their partner, felt they disclosed more information about themselves, felt more comfortable, less awkward, felt more able to relax and be themselves, and also

indicated that if they were to see a therapist they would prefer face-to-face.

3.2 Eye Contact

During face-to-face conversations perceived eye contact was generally high across all interaction types with 61–73% of participants rating there to be eye contact during 70% or greater of the conversation time, see **Figure 6**. Across all interaction types perceived eye contact was significantly lower in VR with 36–44% of participants rating there to be eye contact during 70% or greater of the conversation time, all $ts \geq 3.84$, $ps < 0.001$, $ds \geq 0.58$.

Correlations between the extent of perceived eye contact and the appraisal ratings of the interactions are shown in **Table 2**. Consistent positive associations were found for the ratings of enjoyment, closeness, feeling understood, and comfort. These associations were similar for both face-to-face and virtual reality interactions. This indicates that eye contact is important for the perceived quality of interactions in both modes of communication.

4 DISCUSSION

In this study we explored participant experience in real-time social interaction in virtual reality (VR) where the avatar of the conversational partner was controlled by full face and body motion capture. We refer to such a responsive avatar as a *realistic motion avatar* compared with the semi-realistic motion avatars (with limited motion tracking) of current VR social chat platforms and the more basic avatars (with minimal or no motion tracking) of the past. We contrasted the VR experience with an analogous face-to-face experience across two contexts—Getting acquainted conversation, and a semi-structured interview where participants disclosed negative and positive experiences. As expected, participants had an overall preference for face-to-face interaction compared with interaction in virtual reality. However, as will be discussed, a closer look at our results reveals that the difference in appraisal of the communication modes was only small.

TABLE 2 | Spearman correlations between perceived extent of eye contact with different appraisals of the interaction across different communicative contexts face-to-face and in VR.

	Enjoyment	Closeness	Understood	Disclosure	Comfort	Awkwardness
Getting acquainted						
Face-to-face	0.33*	0.30*	0.24	0.20	0.47*	−0.20
Virtual reality	0.41*	0.26	0.45*	0.32*	0.09	−0.35*
Negative disclosure						
Face-to-face	0.46*	0.36*	0.32*	−0.02	0.31*	−0.20
Virtual reality	0.41*	0.39*	0.32*	−0.08	0.34*	−0.34*
Positive disclosure						
Face-to-face	0.39*	0.36*	0.39*	0.26	0.35*	−0.31*
Virtual reality	0.33*	0.43*	0.35*	0.04	0.34*	−0.19

*$p < 0.05$

The present study provides evidence to suggest that full face and body motion capture of avatars can make social interaction in VR a similar experience to face-to-face interaction. Additionally, some participants reported a preference for interaction in VR, particularly for the context that involved disclosure of negative experiences. We also examined the perception of eye contact across both modes of interaction and found that a high degree of perceived eye contact existed in both modes for most participants. However, we also found the perception of eye contact was diminished in VR for some participants.

4.1 Self-Reported Experience of Conversations Face-to-Face Compared With a Realistic Motion Avatar in VR When Getting Acquainted

In the present study after engaging in face-to-face and VR based getting acquainted conversations participants rated the experiences on several adjectives. No significant difference was found between face-to-face and VR conversations for ratings of enjoyment, feeling understood, self-disclosure, comfort, and awkwardness. Due to our participant's general lack of familiarity with VR we were somewhat surprised that face-to-face was not superior on these ratings. We argue that the naturalness afforded to the interaction by the full facial and body motion tracking of the avatar underlies this result.

The one rating type where face-to-face was rated significantly more positively was for perceived *closeness*. We concede that there may be some ambiguity as to whether participants interpreted the term closeness more in a psychological sense, or in a physical sense. However, there was no significant difference in ratings regarding feeling understood. Therefore, we expect that this difference in perceived closeness has more to do with a sense of physical separation rather than psychological separation. Therefore, like other forms of computer-mediated communication, our results suggest that social interaction in virtual reality may be associated with a greater sense of interpersonal distance compared with face-to-face interaction.

The findings of the current study are consistent with recent research reporting that interactions with motion captured avatars are generally rated positively by participants (Higgins et al., 2021; Seymour et al., 2021; Wu et al., 2021). However, when asking

participants to indicate a preference for either face-to-face or VR we found that face-to-face interaction is preferred by most. This is not surprising considering prior research has found face-to-face interaction is what people are most familiar with and typically preferred over computer-mediated communication (Baltes et al., 2002; Vlahovic et al., 2012; Sprecher, 2014; Sprecher & Hampton, 2017). Despite this, around 10–33% of participants still indicated that in the VR interaction they felt the most comfortable, relaxed and engaged in more self-disclosure. Our results indicate that there is a lot of potential for social interaction in VR with realistic motion avatars to become popular with a substantial number of people. This bodes well for the future platforms for casual social interaction in VR such as Facebook Horizon and others. We predict that head mounted displays of the future will come with eye and mouth motion capture as standard features. Recent examples of headsets incorporating some facial tracking are the *Deca Gear* headset and *HTC Vive* headset (*via* accessories). We anticipate that future technological advancements will enable full motion capture of the face (rather than just eyes and mouth) in headsets of the more distant future.

4.2 Self-Reported Experience of Conversations Face-To-Face Compared With a Realistic Motion Avatar in VR When Disclosing Negative and Positive Experiences

The self-reported ratings made by participants after disclosing negative and positive experiences were largely similar to their prior ratings from the getting acquainted context. That is, there were no significant differences in ratings of self-reported enjoyment, comfort, awkwardness, and self-disclosure between the face-to-face and VR modes. For positive disclosure there was no significant difference for feeling understood, but for negative disclosure overall participants reported feeling more understood when face-to-face. Again, the ratings of closeness were higher interacting face-to-face for both negative and positive disclosure. Overall, in combination with the getting acquainted findings these results indicate that real-time social interaction in VR with a realistic motion avatar has a lot of future potential. More research is needed to investigate the utility of such an

approach as an option for psychological therapy (Baccon et al., 2019; Slater et al., 2019; Pedram et al., 2020; Geraets et al., 2021; Pimentel et al., 2021; Sampaio et al., 2021).

The preference data indicated that during positive disclosure only 10% of participants preferred VR for self-disclosure, comfort, and being able to relax and be oneself. However, for negative disclosure it was approximately 30% preferring VR on the same facets. We tentatively suggest that the increased sense of interpersonal distance in VR may be why some of our participants indicated such preferences for VR over face-to-face for negative disclosure. As suggested by Baccon et al. (2019), social interaction in VR has potential to fill a niche for psychological therapy where there is a little more distance compared with face-to-face, yet greater closeness than other forms of computer-mediated communication such as video chat, phone chat, or text chat. More research is required to understand how interaction with realistic motion avatars in VR compares to other forms of computer-mediated communication for therapeutic purposes (Pedram et al., 2020). While future advancements in technology will likely continue to reduce the discrepancy between face-to-face and VR social interaction, in therapeutic contexts with some clients it may prove advantageous to maintain some extra interpersonal distance in VR interaction.

4.3 Self-Reported Experience of Eye Contact During Conversations Face-To-Face Compared With a Realistic Motion Avatar

Consistent with prior literature in the present study positive associations were found between perceptions of greater levels of eye contact with higher levels of conversational enjoyment, closeness, and comfort (Kleinke, 1986; Akechi et al., 2013; Senju & Johnson, 2009). Additionally, consistent with prior literature perceptions of eye contact were generally very high for face-to-face interaction with 61–73% of participants across the different types of interaction (i.e., getting acquainted, negative disclosure, and positive disclosure) rating eye contact occurring more than 70% of the conversation time (Rogers et al., 2018; Rogers et al., 2019).

During VR interaction eye contact was also rated as generally high with 36–44% of participants rating eye contact as present for more than 70% of the conversation time. However, perceived eye contact in VR was found to be significantly lower compared with face-to-face interactions. Future research is required to further investigate the factors in VR social interaction that act to increase or decrease perceptions of eye contact.

Prior research has indicated that eye contact is associated with increased physiological arousal (Jarick & Bencic, 2019), however it has also been found that eye contact in VR may not elicit the same level of arousal compared with face-to-face eye contact (Syrjamaki et al., 2020). Therefore, in the present study the reduced sense of closeness experienced might be linked to not only the diminished perception of eye contact in VR, but also a lower physiological arousal in VR social interaction even when

eye contact is perceived. This is an intriguing avenue for future research.

4.4 Limitations and Future Research

In the current study the two major differences found between the face-to-face and VR interactions were lower perceptions of closeness and eye contact in VR. A limitation of the study was that the physical distance between the participant and the avatar in VR was slightly greater than the face-to-face conversations. Additionally, the avatar in VR was standing whereas in the face-to-face conditions the conversational partner was sitting. This was done to avoid clipping of the avatar with the table in the VR interactions which might have negatively impacted perceptions of the VR conversation. However, adding the extra distance in VR to avoid the clipping reduces our certainty in the findings for perceived closeness and eye contact. In future research we plan on having the avatar closer to the participant in VR conversations to check if the reduced perceptions of closeness and eye contact were not simply a result of the slight difference in physical distance between participant and conversational partner across face-to-face and VR modes of interaction in the present study.

In their recent study Seymour et al. (2021) argued for a superiority of photo-realistic avatars over avatars of lower graphical fidelity. A limitation of our study was that while the avatars used were modelled to look like the actual people the represented, they were not of a photo-realistic quality. Seymour et al. (2021) did not use full body motion capture for any of the avatars in their study. In our study, the avatars may not have been photo-realistic, yet graphical fidelity was still relatively high, and importantly the behavioural realism was high because of the full face and body motion tracking techniques implemented. Further research is required to tease apart the roles that both graphical realism and behavioural realism play in shaping people's experiences with social interaction in VR (Pan & Hamilton, 2018; Seymour et al., 2018; Zibrek et al., 2019; Ferstl et al., 2021; Seymour et al., 2021; Zibrek et al., 2021). We tentatively suggest that behavioural realism might be the more important factor for influencing conversational enjoyment in VR.

In our study we compared social interaction with realistic motion avatars to face-to-face conversation. Our conclusions are therefore limited to such a comparison. More research is required to compare social interaction in HMD-based virtual reality with realistic motion avatars with other technology-mediated modes of interaction such as text/email messages, phone chat, and video chat (Higgins et al., 2021; Seymour et al., 2021).

Another limitation of the present study was that while the participants were in VR, the person they were interacting with controlling the motion captured avatar was not. Furthermore, we did not provide an avatar for the participant which is an important element to enrich the experience (Latoschik et al., 2017; Pan & Steed, 2017; Freeman et al., 2020). As we have mentioned, new HMDs are on the horizon with integrated facial motion tracking that will enable future investigations where both interactants are fully motion captured while both parties are in VR. Regardless, there might be certain contexts where the method used in the current study (i.e., one person in VR while the other is not) might be advantageous. For example, it may be advantageous

for certain therapist—client interactions where the client is in VR while the therapist is not so that the therapist can be observing the real-world body language of the client. It may also help to limit fatigue of therapists if they are interviewing multiple clients in VR across extended periods of time. Future research is needed to investigate such issues of client-therapist interactions in VR using realistic motion avatars.

The sample in the present study was predominately 30-year-old socially active females with limited experience with virtual reality. Research suggests that repeated exposure to computer mediated communication can act to enhance the perception of richness of that mode (Fernandez et al., 2013; Khojasteh & Won, 2021). Furthermore, we expect that other types of participants might be more open-minded about the technology. Therefore, our results might provide an underestimation of people's ratings and preferences for VR interaction. For example, socially anxious people might respond well to the heightened sense of interpersonal distance that VR interaction could provide (Shalom et al., 2015; Kroczek et al., 2020). Also, younger people that have grown up with greater exposure to digital communication technologies and computer games are likely to be more open-minded about the possibilities of social interaction in virtual spaces (Center, 2015; Lenhart et al., 2015). People with limited social interaction who yearn for more social contact might benefit from VR interactions (Gentina & Chen, 2019; Liddle et al., 2020; Thach et al., 2020). Especially if those VR interactions can incorporate haptic devices to produce sensations that simulate physical contact to complement the social experience (Cui et al., 2021).

Online video chat platforms can be experienced as awkward as this type of computer-mediated communication produces an experience of excessive eye contact and diminished body language that can negatively impact upon turn-taking dynamics and disrupt the flow of conversation, especially in group chat (Bailenson, 2021). Therefore, VR meetings with realistic avatars with full face and body motion tracking has potential to facilitate the productivity of such interactions (Wu et al., 2021). Enhanced communication *via* realistic motion avatars can potentially enhance socially shared joint activities in virtual worlds. We therefore argue that the use of realistic motion avatars will facilitate the popularity and development of virtual tourism experiences (Mura et al., 2017; Hudson et al., 2019). In educational contexts the use of realistic motion avatars has the potential to enhance learning virtual experiences across all levels (i.e., primary, secondary, and higher education) (Liang et al., 2016; Hu-Au & Lee, 2017; Kavanagh et al., 2017; de Siqueira et al., 2021; Hamilton et al., 2021). There is clearly a wide range of contexts where realistic avatars have the potential to enhance virtual experiences and there are many avenues for future research in this area. We have alluded to just a few such avenues here.

REFERENCES

Akechi, H., Senju, A., Uibo, H., Kikuchi, Y., Hasegawa, T., and Hietanen, J. K. (2013). Attention to Eye Contact in the West and East: Autonomic Responses and Evaluative Ratings. *PLoS ONE* 8 (3), e59312. doi:10.1371/journal.pone.0059312

5 CONCLUSION

In the present study we demonstrate how people interacting with a realistic motion avatar rate the experience as fairly similar to face-to-face interaction. We demonstrated this across two conversational contexts—Getting acquainted conversation, and structured interviews designed to elicit participant self-disclosure about negative and positive events. We suggest that harnessing motion capture to enhance social interaction in VR will catapult this mode of computer mediated communication to become the next major evolution in communication technology. Future research is required to determine how to best harness the technology across different communicative contexts, such as casual conversation, therapeutic settings, business settings, tourist settings, education settings, among others.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: Figshare—https://doi.org/10.6084/m9.figshare.15085749.v1.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Edith Cowan University ethics committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SR led the conceptualisation of the study in collaboration with RB, JB, AF, and CS. Data collection was conducted by RB, JB, and AF under supervision of SR. SR conducted data analysis and presentation of data. SR led the drafting of the manuscript in collaboration with CS. All authors read, provided feedback and approved the final manuscript.

FUNDING

Attardo, S., Eisterhold, J., Hay, J., and Poggi, I. (2003). Multimodal Markers of Irony and Sarcasm. *Humor* 16 (2), 243–260. doi:10.1515/humr.2003.012

Baccon, L. A., Chiarovano, E., and MacDougall, H. G. (2019). Virtual Reality for Teletherapy: Avatars May Combine the Benefits of Face-To-Face Communication with the Anonymity of Online Text-Based

Communication. *Cyberpsychology, Behav. Soc. Networking* 22 (2), 158–165. doi:10.1089/cyber.2018.0247

Bailenson, J. N. (2021). Nonverbal Overload: A Theoretical Argument for the Causes of Zoom Fatigue. *Technol. Mind, Behav.* 2 (1). doi:10.1037/tmb0000030

Baltes, B. B., Dickson, M. W., Sherman, M. P., Bauer, C. C., and LaGanke, J. S. (2002). Computer-mediated Communication and Group Decision Making: A Meta-Analysis. *Organizational Behav. Hum. Decis. Process.* 87 (1), 156–179. doi:10.1006/obhd.2001.2961

Baron-Cohen, S., Jolliffe, T., Mortimore, C., and Robertson, M. (1997). Another Advanced Test of Theory of Mind: Evidence from Very High Functioning Adults with Autism or Asperger Syndrome. *J. Child Psychol. Psychiatry* 38, 813–822. doi:10.1111/j.1469-7610.1997.tb01599.x

Center, P. R. (2015). Pewinternet. Available at: http://www.pewinternet.org/2015/08/06/teens-technology-and-friendships/.

Cui, D., Kao, D., and Mousas, C. (2021). Toward Understanding Embodied Human-virtual Character Interaction through Virtual and Tactile Hugging. *Comput. Anim. Virtual Worlds* 32, cav2009. doi:10.1002/cav.2009

Falconer, C. J., Davies, E. B., Grist, R., and Stallard, P. (2019). Innovations in Practice: Avatar-based Virtual Reality in CAMHS Talking Therapy: Two Exploratory Case Studies. *Child. Adolesc. Ment. Health* 24 (3), 283–287. doi:10.1111/camh.12326

Fernandez, V., Simo, P., Sallan, J. M., and Enache, M. (2013). Evolution of Online Discussion Forum Richness According to Channel Expansion Theory: A Longitudinal Panel Data Analysis. *Comput. Education* 62, 32–40. doi:10.1016/j.compedu.2012.10.020

Ferstl, Y., Thomas, S., Guiard, C., Ennis, C., and McDonnell, R. (2021). "Human or Robot? Investigating Voice, Appearance and Gesture Motion Realism of Conversational Social Agents," in Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents, September 2021, Virtual event Japan, 76–83. doi:10.1145/3472306.3478338

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. Thousand Oaks, California, United States: Sage.

Freeman, G., Zamanifard, S., Maloney, D., and Adkins, A. (2020). My Body, My Avatar: How People Perceive Their Avatars in Social Virtual Reality. *Extended Abstr. 2020 CHI Conf. Hum. Factors Comput. Syst.*, 1–8. doi:10.1145/3334480.3382923

Gentina, E., and Chen, R. (2019). Digital Natives' Coping with Loneliness: Facebook or Face-To-Face? *Inf. Management* 56, 103138. doi:10.1016/j.im.2018.12.006

Geraets, C. N. W., van der Stouwe, E. C. D., Pot-Kolder, R., and Veling, W. (2021). Advances in Immersive Virtual Reality Interventions for Mental Disorders: A New Reality? *Curr. Opin. Psychol.* 41, 40–45. doi:10.1016/j.copsyc.2021.02.004

Gomes de Siqueira, A., Feijóo-García, P. G., Stuart, J., and Lok, B. (2021). Toward Facilitating Team Formation and Communication through Avatar Based Interaction in Desktop-Based Immersive Virtual Environments. *Front. Virtual Real.* 2. doi:10.3389/frvir.2021.647801

Gunkel, S., Stokking, H., Prins, M., Niamut, O., Siahaan, E., and Cesar, P. (2018). "Experiencing Virtual Reality Together," in Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video, June 2018, 233–238. doi:10.1145/3210825.3213566

Hamilton, D., McKechnie, J., Edgerton, E., and Wilson, C. (2021). Immersive Virtual Reality as a Pedagogical Tool in Education: A Systematic Literature Review of Quantitative Learning Outcomes and Experimental Design. *J. Comput. Educ.* 8, 1–32. doi:10.1007/s40692-020-00169-2

Higgins, D., Fribourg, R., and McDonnell, R. (2021). Remotely Perceived: Investigating the Influence of Valence on Self-Perception and Social Experience for Dyadic Video-Conferencing with Personalized Avatars. *Front. Virtual Real.* 2, 668499. doi:10.3389/frvir.2021.668499

Ho, S., Foulsham, T., and Kingstone, A. (2015). Speaking and Listening with the Eyes: Gaze Signaling during Dyadic Interactions. *PLoS ONE* 10, e0136905. doi:10.1371/journal.pone.0136905

Hu Au, E., and Lee, J. J. (2017). Virtual Reality in Education: A Tool for Learning in the Experience Age. *Ijiie* 4 (4), 215–226. doi:10.1504/IJIIE.2017.09148110.1504/ijiie.2017.10012691

Hudson, S., Matson-Barkat, S., Pallamin, N., and Jegou, G. (2019). With or without You? Interaction and Immersion in a Virtual Reality Experience. *J. Business Res.* 100, 459–468. doi:10.1016/j.jbusres.2018.10.062

Jarick, M., and Bencic, R. (2019). Eye Contact Is a Two-Way Street: Arousal Is Elicited by the Sending and Receiving of Eye Gaze Information. *Front. Psychol.* 10, 1262. doi:10.3389/fpsyg.2019.01262

Karaaslan, A., Durmuş, B., and Amado, S. (2020). Does Body Context Affect Facial Emotion Perception and Eliminate Emotional Ambiguity without Visual Awareness? *Vis. Cogn.* 28 (10), 605–620. doi:10.1080/13506285.2020.1846649

Kavanagh, S., Luxton-Reilly, A., Wuensche, B., and Plimmer, B. (2017). A Systematic Review of Virtual Reality in Education. *Themesin Sci. Technology Education* 10 (2), 85–119. Available at: https://www.learntechlib.org/p/182115/.

Khojasteh, N., and Won, A. S. (2021). Working Together on Diverse Tasks: A Longitudinal Study on Individual Workload, Presence and Emotional Recognition in Collaborative Virtual Environments. *Front. Virtual Real.* 2. doi:10.3389/frvir.2021.643331

Kleinke, C. L. (1986). Gaze and Eye Contact: A Research Review. *Psychol. Bull.* 100 (1), 78–100. doi:10.1037/0033-2909.100.1.78

Kock, N. (2005). Media Richness or media Naturalness? the Evolution of Our Biological Communication Apparatus and its Influence on Our Behavior toward E-Communication Tools. *IEEE Trans. Profess. Commun.* 48 (2), 117–130. doi:10.1109/TPC.2005.849649

Kroczek, L. O. H., Pfaller, M., Lange, B., Müller, M., and Mühlberger, A. (2020). Interpersonal Distance during Real-Time Social Interaction: Insights from Subjective Experience, Behavior, and Physiology. *Front. Psychiatry* 11, 561. doi:10.3389/fpsyt.2020.00561

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 Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Tachnology, November 2017, Gottenburg, Sweden, 1–10. doi:10.1145/3139131.3139156

Lenhart, A., Smith, A., Anderson, M., Duggan, M., and Perrin, A. (2015). *Teens, Technology & Friendships: Video Games, Social media and mobile Phones Play an Integral Role in How Teens Meet and Interact with Friends*. Washington, DC: Pew Internet.

Liang, H., Chang, J., Deng, S., Chen, C., Tong, R., and Zhang, J. J. (2016). Exploitation of Multiplayer Interaction and Development of Virtual Puppetry Storytelling Using Gesture Control and Stereoscopic Devices. *Comput. Anim. Virtual Worlds* 28, e1727. doi:10.1002/cav.1727

Liddle, J., Stuart, A., Worthy, P., Levine, M., Kastelle, T., Wiles, J., et al. (2020). "Building the Threads of Connection that We Already Have": The Nature of Connections via Technology for Older People. *Clin. Gerontologist* 44, 406–417. doi:10.1080/07317115.2020.1852638

Maloney, D., Freeman, G., and Wohn, D. Y. (2020). "Talking without a Voice". *Proc. ACM Hum.-Comput. Interact.* 4 (175), 1–25. doi:10.1145/3415246

Meeren, H. K. M., van Heijnsbergen, C. C. R. J., and de Gelder, B. (2005). Rapid Perceptual Integration of Facial Expression and Emotional Body Language. *Proc. Natl. Acad. Sci.* 102 (45), 16518–16523. doi:10.1073/pnas.0507650102

Meier, J. V., Noel, J. A., and Kaspar, K. (2021). Alone Together: Computer-Mediated Communication in Leisure Time during and after the COVID-19 Pandemic. *Front. Psychol.* 12. doi:10.3389/fpsyg.2021.666655

Mura, P., Tavakoli, R., and Pahlevan Sharif, S. (2017). 'Authentic but Not Too Much': Exploring Perceptions of Authenticity of Virtual Tourism. *Inf. Technol. Tourism* 17 (2), 145–159. doi:10.1007/s40558-016-0059-y

Pan, X., and Hamilton, A. F. D. C. (2018). Why and How to Use Virtual Reality to Study Human Social Interaction: The Challenges of Exploring a New Research Landscape. *Br. J. Psychol.* 109, 395–417. doi:10.1111/bjop.12290

Pan, Y., and Steed, A. (2017). The Impact of Self-Avatars on Trust and Collaboration in Shared Virtual Environments. *PLoS ONE* 12 (12), e0189078. doi:10.1371/journal.pone.0189078

Pedram, S., Palmisano, S., Perez, P., Mursic, R., and Farrelly, M. (2020). Examining the Potential of Virtual Reality to Deliver Remote Rehabilitation. *Comput. Hum. Behav.* 105, 106223. doi:10.1016/j.chb.2019.106223

Pettersson, E., and Sundstedt, V. (2017). A Perceptual Evaluation of Social Interaction with Emotes and Real-Time Facial Motion Capture. *Proc. Tenth Int. Conf. Motion Games*, 1–6. doi:10.1145/3136457.3136461

Pfund, G. N., Harriger, J., and Hill, P. L. (2021). Video Chat Usage and the Big Five in Women during the COVID-19 Pandemic. *Personal. Individual Differences* 171, 110537. doi:10.1016/j.paid.2020.110537

Pimentel, D., Foxman, M., Davis, D. Z., and Markowitz, D. M. (2021). Virtually Real, but Not Quite There: Social and Economic Barriers to Meeting Virtual

Reality's True Potential for Mental Health. *Front. Virtual Real.* 2, 627059. doi:10.3389/frvir.2021.627059

Rehm, I. C., Foenander, E., Wallace, K., Abbott, J.-A. M., Kyrios, M., and Thomas, N. (2016). What Role Can Avatars Play in E-Mental Health Interventions? Exploring New Models of Client-Therapist Interaction. *Front. Psychiatry* 7, 186. doi:10.3389/fpsyt.2016.00186

Rogers, S. L., and Cruickshank, T. (2021). Change in Mental Health, Physical Health, and Social Relationships during Highly Restrictive Lockdown in the COVID-19 Pandemic: Evidence from Australia. *PeerJ* 9, e11767. doi:10.7717/peerj.11767

Rogers, S. L., Guidetti, O., Speelman, C. P., Longmuir, M., and Phillips, R. (2019). Contact Is in the Eye of the Beholder: The Eye Contact Illusion. *Perception* 48 (3), 248–252. doi:10.1177/0301006619827486

Rogers, S. L., Speelman, C. P., Guidetti, O., and Longmuir, M. (2018). Using Dual Eye Tracking to Uncover Personal Gaze Patterns during Social Interaction. *Sci. Rep.* 8, 4271. doi:10.1038/s41598-018-22726-7

Sampaio, M., Navarro Haro, M. V., De Sousa, B., Vieira Melo, W., and Hoffman, H. G. (2021). Therapists Make the Switch to Telepsychology to Safely Continue Treating Their Patients during the COVID-19 Pandemic. Virtual Reality Telepsychology May Be Next. *Front. Virtual Real.* 1, 576421. doi:10.3389/frvir.2020.576421

Senju, A., and Johnson, M. H. (2009). The Eye Contact Effect: Mechanisms and Development. *Trends Cogn. Sci.* 13 (3), 127–134. doi:10.1016/j.tics.2008.11.009

Seymour, M., Riemer, K., Riemer, K., and Kay, J. (2018). Actors, Avatars and Agents: Potentials and Implications of Natural Face Technology for the Creation of Realistic Visual Presence. *Jais* 19 (10), 953–981. doi:10.17705/1jais.00515

Seymour, M., Yuan, L., Dennis, A. R., and Riemer, K. (2021). Have We Crossed the Uncanny valley? Understanding Affinity, Trsutworthiness, and Preference for Realistic Digital Humans in Immersive Environments. *J. Assoc. Inf. Syst.* 22 (3), 591–617. doi:10.17705/1jais.00674

Shalom, J. G., Israeli, H., Markovitzky, O., and Lipsitz, J. D. (2015). Social Anxiety and Physiological Arousal during Computer Mediated vs. Face to Face Communication. *Comput. Hum. Behav.* 44, 202–208. doi:10.1016/j.chb.2014.11.056

Slater, M., Neyret, S., Johnston, T., Iruretagoyena, G., Crespo, M. Á. d. l. C., Alabèrnia-Segura, M., et al. (2019). An Experimental Study of a Virtual Reality Counselling Paradigm Using Embodied Self-Dialogue. *Sci. Rep.* 9 (1), 1–13. doi:10.1038/s41598-019-46877-3

Smith, H. J., and Neff, M. (2018). "Communication Behavior in Embodied Virtual Reality," in Proceedings of the 2018 CHI conference on human factors in computing systems, April 2018, Montreal QC Canada, 1–12. doi:10.1145/3173574.3173863

Sprecher, S. (2021). Closeness and Other Affiliative Outcomes Generated from the Fast Friends Procedure: A Comparison with a Small-Talk Task and Unstructured Self-Disclosure and the Moderating Role of Mode of Communication. *J. Soc. Personal Relationships* 38 (5), 1452–1471. doi:10.1177/0265407521996055

Sprecher, S., Hampton, A. J., Heinzel, H. J., and Felmlee, D. (2016). Can I Connect with Both You and My Social Network? Access to Network-Salient Communication Technology and Get-Acquainted Interactions. *Comput. Hum. Behav.* 62, 423–432. doi:10.1016/j.chb.2016.03.090

Sprecher, S., and Hampton, A. J. (2017). Liking and Other Reactions after a Get-Acquainted Interaction: A Comparison of Continuous Face-To-Face Interaction versus Interaction that Progresses from Text Messages to Face-To-Face. *Commun. Q.* 65 (3), 333–353. doi:10.1080/01463373.2016.1256334

Sprecher, S. (2014). Initial Interactions Online-Text, Online-Audio, Online-Video, or Face-To-Face: Effects of Modality on Liking, Closeness, and Other Interpersonal Outcomes. *Comput. Hum. Behav.* 31, 190–197. doi:10.1016/j.chb.2013.10.029

Suler, J. R. (2000). Psychotherapy in Cyberspace: A 5-Dimensional Model of Online and Computer-Mediated Psychotherapy. *CyberPsychology Behav.* 3 (2), 151–159. doi:10.1089/109493100315996

Syrjämäki, A. H., Isokoski, P., Surakka, V., Pasanen, T. P., and Hietanen, J. K. (2020). Eye Contact in Virtual Reality - A Psychophysiological Study. *Comput. Hum. Behav.* 112, 106454. doi:10.1016/j.chb.2020.106454

Thach, K. S., Lederman, R., and Waycott, J. (2020). How Older Adults Respond to the Use of Virtual Reality for Enrichment: A Systematic Review. *OZCHI* 20, 303–313. doi:10.1145/3441000.3441003

Vesper, C., and Sevdalis, V. (2020). Informing, Coordinating, and Performing: A Perspective on Functions of Sensorimotor Communication. *Front. Hum. Neurosci.* 14, 168. doi:10.3389/fnhum.2020.00168

Vlahovic, T. A., Roberts, S., and Dunbar, R. (2012). Effects of Duration and Laughter on Subjective Happiness within Different Modes of Communication. *J. Comput-mediat Comm.* 17, 436–450. doi:10.1111/j.1083-6101.2012.01584.x

Wu, Y., Wang, Y., Jung, S., Hoermann, S., and Lindeman, R. W. (2019). Exploring the Use of a Robust Depth-Sensor-Based Avatar Control System and its Effects on Communication Behaviors. *Proc. 25th ACM Symp. Virtual Reality Softw. Technology*, 1–9. doi:10.1145/3359996.3364267

Wu, Y., Wang, Y., Jung, S., Hoermann, S., and Lindeman, R. W. (2021). Using a Fully Expressive Avatar to Collaborate in Virtual Reality: Evaluation of Task Performance, Presence, and Attraction. *Front. Virtual Real.* 2, 641296. doi:10.3389/frvir.2021.641296

Zibrek, K., Cabral, J. P., and McDonnell, R. (2021). Does Synthetic Voice Alter Social Response to a Photorealistic Character in Virtual Reality? *Motion, Interaction and Games* 21, 1–6. doi:10.1145/3487983.3488296

Zibrek, K., Martin, S., and McDonnell, R. (2019). Is Photorealism Important for Perception of Expressive Virtual Humans in Virtual Reality? *ACM Trans. Appl. Perception*, 16, 1–19. doi:10.1145/3349609