Self-reported sense of presence and responses to social stimuli by adolescents with ASD in a collaborative virtual reality environment

Simon Wallace[1], Sarah Parsons[2] and Anthony Bailey[3]

[1] Division of Clinical Neurology, Level 6, West Wing, John Radcliffe Hospital, Oxford OX3 9DU

[2] Southampton Education School, University of Southampton, Building 32, Southampton, SO17 1BJ, United Kingdom

[3] UBC Department of Psychiatry, 2255 Westbrook Mall, Vancouver, V6T 2A1, Canada

*Author for correspondence: S.J.Parsons@soton.ac.uk

Author Note

This research was conducted without any external funding. The authors have no conflicts of interest in relation to this study.

Abstract

Background Collaborative Virtual Environments (CVEs) have the potential to support socio-communicative interactions for people with autism spectrum disorder (ASD), but little is known about the sense of presence participants feel in CVEs or how CVEs can be used to assess skills.

Method Ten children with ASD and 10 typically developing (TD) children (aged 12-16 years) judged greeting behaviours of a human avatar and static facial expressions in a virtual gallery.

Results There were no differences in presence reported by the two groups. The ASD group were less sensitive to a negative greeting from the human avatar than the TD group, and impaired in recognising static facial expressions.

Conclusions Self-reported measures of presence are valuable for informing which kinds of tasks and technology may provide more authentic contexts in which to identify and support social competence in participants with ASD.

Introduction

Since the mid-1990s researchers have investigated how virtual reality (VR) technologies can support the learning of individuals with an autism spectrum disorder (ASD; e.g., Strickland Marcus, Mesibov, & Hogan, 1996; Trepagnier, 1999). A particular focus has been the application of VR to supporting and practising social skills in motivating and minimally threatening environments that can be carefully designed and controlled (Parsons & Mitchell, 2002) while also supporting more lifelike and less didactic interactions (Kandalaft, Didehbani, Krawczyk, Allen, & Chapman, 2013). VR technology has relevance to the disability field generally because of its capacity to help overcome the physical and cognitive barriers to social inclusion (Stendal, Balandin, & Molka-Danielsen, 2011). In addition, such technologies may offer a cost-effective and accessible means of assessing, and targeting, children's learning needs (Goodwin, 2008; Grynszpan, Weiss, Perez-Diaz, & Gal, 2014; Knight, McKissick, & Saunders, 2013).

In a review of the ASD and VR field, Parsons and Cobb (2011) reported mostly positive messages for children and young people with ASD: Children with ASD can use VR technologies successfully (Charitos et al., 2000; Parsons, Mitchell, & Leonard, 2004); they do not find them aversive (Wallace et al., 2010); and they interpret virtual environments (VEs) where virtual characters appear on screen, as representing aspects of the real world through making judgements about the similarity between real and virtual scenes (Parsons et al., 2004; Parsons, Mitchell, & Leonard, 2005). In addition, studies have shown that children with ASD can apply information learned within VEs to real world settings (Josman, Milika, Ben-Chaim, Friedrich, & Weiss, 2008; Strickland, McAllister, Coles, & Osborne, 2007). Nevertheless, Parsons and Cobb (2011) found research in this area to be limited overall, and highlighted a need to "…understand how best to use the technology and develop our understanding about how to construct VEs that are meaningful and applicable to the learning needs of users" (p.362). Similarly, in a meta-analysis of the effectiveness of innovative technologies for ASD, Grynszpan et al. (2014) concluded that although there were some indications that VR could be effective as a learning intervention "…there is still little evidence supporting the efficacy of this technology, and more studies investigating this promising area of research are called for" (p.358).

Interpretation and responding in virtual scenes: measuring presence

A central assumption guiding the use of VEs is that they create an authentic environment in which socio-communicative skills can be learned and practised (Parsons & Mitchell, 2002). This assumption requires that children with ASD experience VEs as being realistic and truthful (Blascovich et al., 2002). Otherwise, if VEs seemed too "game-like" and suspended from reality the aim of supporting real world skills and understanding would seem doubtful. Moreover, for VEs to have clinical and educational benefit for supporting real world understanding then there needs to be some similarity between responses in VEs and responses in the real world in order that generalization of experiences across contexts may be more likely to occur (Kandalaft et al., 2013; Parsons et al., 2005; Schroeder, 2002). As Yee, Bailenson, Urbanek, Chang and Merget (2007) argue: "…to understand how large the potential a medium has to change an individual, researchers have typically measured how realistically a user behaves while inside of that medium" (p.116).

In VR research of individuals without ASD there has been a long tradition of establishing how realistic and truthful experiences in VEs are by measuring the core psychological concept of "presence" , that is: "…the sense of being caught up in the representations of virtual worlds" (Jacobson, 2001, p.653). Although the concept (and its measurement) are debated (e.g., Slater, 2004) there is, nevertheless, broad agreement that the degree of presence experienced in VEs is related to the potential of the medium for supporting learning (Mikropoulos, 2006; Mikropoulos & Natsis, 2011). Thus, understanding the circumstances under which participants feel present and engaged in VEs is an important line of enquiry. Such research has revealed the influence of various features on feelings of presence and responding in VR, including individual characteristics of users without ASD, such as personality (e.g., Hammick & Lee, 2014) and spatial ability (Waller, 2000); and the different features of the technology, such as how anthropomorphic the virtual characters are within the VEs (e.g., Nowak, 2004; Nowak & Biocca, 2003).

Investigating presence in ASD

Despite the well-established importance of understanding presence in VR research, it is very surprising that only one study has thus far examined the concept directly with participants with ASD. Wallace et al. (2010) investigated the experiences of adolescents with ASD and typically developing (TD) peers in graphically rich and sophisticated VEs, which did not require a head-mounted device. The two groups experienced and interpreted the scenes in a similar way, that is, they reported similar feelings of presence as measured by the Independent Television Commission – Sense of Presence Inventory questionnaire (ITC SoPI; Lessiter, Freeman, Keogh, & Davidoff, 2001). Also, there were no differences between the groups in reported negative sensations, in contrast to what might have been predicted given the sensory and perceptual experiences of some people with ASD (Leekam, Nieto, Libby, Wing, & Gould, 2007; Mottron, Dawson, Soulieres, Hubert, & Burack, 2006). This finding has been interpreted as showing that ASD and TD participants experience VEs similarly and, thus, VEs "…can reliably be used to simulate authentic social situations in experimental settings" (Georgescu, Kuzmanovic, Roth, Bente, & Vogeley, 2014, p.7).

Indeed, despite these similarities between the ASD and TD groups in the study of Wallace et al. (2010), the groups differed when rating their "social attraction" (Nowak, 2004) towards two different avatars in a VE. The TD group rated a friendly avatar (a computerised child character called Danny who was presented in a virtual playground) as more socially attractive (e.g., likeable; would make a good friend) than an unfriendly character, but the ASD group rated the two characters similarly. Wallace et al. (2010) suggested that the general sense of immersion of the ASD group in the VE was typical and that their difficulty in distinguishing the social intent of virtual characters reflected their real- world social impairment (Baron-Cohen, 2000). Nevertheless, difficulties in interpreting the motivation of the virtual characters (avatars) could be because they were preprogrammed and under the control of the computer (agent avatars) rather than dynamic and responsive (under the control of a teacher or therapist; i.e., human-controlled avatars). There is evidence that children with ASD, like TD individuals (Blascovich et al., 2002), are sensitive to whether virtual characters are real or not and report that this knowledge influences their responses in VEs (Parsons et al., 2005; Parsons, Mitchell, & Leonard, 2006). It is also well established that the nature of stimuli used in prompting social responses by participants with ASD influences their judgements, and these judgements also differ from TD individuals. For example, participants with ASD engage differently with “synthetic” characters depending on whether the characters are photographic, cartoon-like, 3D-avatar like, dynamic, or static (e.g., Forgeot d’Arc et al., 2014; Riby & Hancock, 2009; Rosset et al., 2008). This suggests that it is important to establish whether a similar pattern of results to Wallace et al. (2010) would be found in response to human-controlled avatars rather than computer-controlled agent avatars.

Responses of participants with ASD in Collaborative Virtual Environments

The technology used in Wallace et al. (2010) was a proof-of-concept large immersive VE, unlikely to bear much relevance to real world contexts of use in, for example, schools and clinics. Consequently, it is important to examine participants’ reports of presence when interacting with a human-controlled avatar in a desktop collaborative VE (CVE), which supports the concurrent interactions of two or more users via the internet (similar to computer games where two players interact on screen). Such a set-up is much more likely to inform technical configurations and interactions for supporting eventual assessment and learning using VR technologies than the VEs used in Wallace et al. (2010). Specifically, given the substantial costs and practical considerations that are often involved in assessing children with ASD, and in taking part in research, it is important to understand whether a desktop CVE offers an ecologically valid context in which to investigate and support social responding. Given that desktop (or laptop) CVEs utilise technology hardware that many schools, families and clinics have access to, they may obviate the need to bring teachers, therapists, parents and children together in the same place, and thus offer considerable potential for distance learning and assessment (Stichter, Laffey, Galyen, & Herzog, 2014).

Recognising this potential, there is a small but growing interest in the application of CVEs for supporting learning for participants with ASD. For example, research has incorporated avatars into desktop environments to support the learning of vocabulary (Bosseler & Massaro, 2003; Massaro & Bosseler, 2006) and to investigate the interpretation of facial expressions (Fabri, Elzouki, & Moore, 2007; Fabri & Moore, 2005; Grynszpan et al., 2009; Moore, Cheng, McGrath, & Powell, 2005; Schwartz, Bente, Gawronski, Schilbach, & Vogeley, 2010). However, these studies used prerecorded agent avatars rather than human-controlled avatars to prompt responding and so tell us very little about how participants with ASD respond to human-controlled avatars in real time. Other studies using desktop or laptop CVEs have investigated the interactions of participants with ASD with other virtual characters played by teachers (Cheng & Ye, 2010), students and teachers (Schmidt, Laffey, Schmidt, Wang, & Stichter, 2012; Stichter et al., 2014) or a clinician (Kandalaft et al., 2013). These studies provide preliminary evidence that students and adults with ASD do respond socially to virtual characters in the VE (through supporting conversations, use of gestures and facial expressions), suggesting that these environments could offer useful platforms for intervention.

However, although Kandalaft et al. (2013) provided anecdotal evidence from their adult participants about the impact the participants felt the VEs had had on their social skills and cognition, the participants' sense of presence and engagement with the VEs were not measured directly. Stichter et al. (2014) gained parents' and teachers' judgements about the social validity of their distance learning CVE scenarios but did not ask the students themselves for their views; consequently, we do not know the extent to which the participants were engaged in the tasks and felt them to be realistic. Such information could be helpful for interpreting, at least in part, the high variability in responding, and lack of any independently-observed learning effects, of the participants in that study. Additionally, the studies previously mentioned did not include any TD participants or provide any information regarding the responses of TD participants to the stimuli presented. Although there is value in investigating the responses of participants with ASD in VEs, a lack of comparison groups means that it is difficult to determine to what extent responding in a VE is related to ASD-specific characteristics or to particular technical or representational characteristics of the VE (cf. Ke & Im, 2013; Nowak & Biocca, 2003; Schmidt et al., 2012).

The present study

In the present study we investigated the self-reported sense of presence of a group of adolescents with ASD to brief social scenarios experienced in a desktop CVE and compared these with an age-matched TD group. Social judgements about these scenarios were explored in two ways: first, in a social attraction task similar to that used in Wallace et al. (2010) to assess whether participants differentiated between positive and negative human-controlled avatar initiated interactions within a CVE, and second, in a recognition task of static facial expression stimuli displayed in a virtual gallery. The latter task was included because it is a measure that shows reliable difficulties in recognition of facial expressions fear, sad and disgust (but not surprise or angry) amongst participants with ASD in computer-based experimental testing (Wallace, Coleman, & Bailey, 2008), and thus can be used as a validity check for performing this task within this particular CVE set-up. In summary, our research questions were:

(1) Relative to age and IQ-matched adolescents with typical development, do adolescents with ASD self-report the same level of presence or involvement in a desktop CVE?; and

(2) Can difficulties associated with understanding social greetings and recognising facial expressions attributed to adolescents with ASD in real world settings be replicated in a desktop CVE?

Method

Design

This was a group-based comparison study in which participants with and without an ASD individually took part in a series of tasks in a desktop CVE during a single session that lasted approximately 40 minutes. The session took place in a quiet room within a children's hospital building in Oxfordshire, UK. During the session, participants completed specific tasks requiring social judgements and were asked to complete structured questionnaires about their sense of presence and feelings of social attraction towards the human avatar. The tasks proceeded in the order described in the procedure section.

Participants

Twenty participants were recruited, comprising an ASD group and a typically developing (TD) control group (with no family history of ASD). The ASD group comprised 10 adolescents (nine male) aged 12 – 16 years (M = 14.8, SD = 1.1). All participants with ASD had a clinical diagnosis (two diagnosed with autism, eight with Asperger syndrome) provided by an independent clinician according to ICD-10 criteria (World Health Organization, 1992) and were recruited through a local school in the south of England. The participants with ASD all attended an ASD specialist unit attached to a mainstream (general) school, where a formal diagnosis is required for entry. The TD participants attended the mainstream school to which the specialist ASD unit was attached. The Autism Diagnostic Interview-Revised (ADI–R; Lord, Rutter, & Le Couteur, 1994) was used by the first author to confirm these diagnoses; most participants scored above the cut-off on the social, communication, repetitive and developmental delay domains, except for two participants (one individual scored one point below and another two points below on the repetitive domain). The TD group comprised 10 adolescents (eight male) aged 14 – 16 years (M = 15.4, SD = .8).

There was no formal matching of participants beyond recruiting from similar age groups and trying to achieve a similar balance for gender. Given the average intelligence profiles of the ASD group, the likelihood that TD volunteers would match on overall IQ profile was high. Indeed, there were no significant group differences in chronological age or verbal, performance or full-scale IQ (assessed by the Wechsler Abbreviated Scale of Intelligence; Wechsler, 1999). Table 1 presents a breakdown of background characteristics of the two groups. Written assent was collected from the young people involved and written informed consent was obtained from parents prior to their child’s participation. The research was reviewed by Oxfordshire NHS Research Ethics Committee and full approval was received prior to the commencement of the study. The main ethical issue highlighted by the committee was the clarity of the information sheets, particularly for the children. Advice was sought from colleagues and local teachers before revisions were made and all participants were fully informed and had the chance to ask questions prior to starting the study.

Insert TABLE 1 about here

Measures

Sense of Presence: The Independent Television Commission – Sense of Presence Inventory (ITC – SoPI; Lessiter et al., 2001) is a 44-item self-rated questionnaire used to measure how involved or present participants feel when experiencing different media. It is a widely used measure of presence within the literature, demonstrating good reliability and validity, that provides 5-point Likert scale data (from *strongly agree* to *strongly disagree*) on four main factors: (a) an individual's level of Spatial Presence: 19 items (e.g., It felt like I was visiting the places in the virtual environment; I felt I wasn't just watching something, like TV) (b) their Engagement with the content: 13 items (e.g., I would have liked the experience to continue; I felt myself being drawn into the virtual environment); (c) the Ecological Validity / Naturalness of the content: 5 items (e.g., The virtual environment seemed like real life; The virtual environment seemed believable to me); and (d) whether there were any Negative Effects from experiencing the VE: 6 items (e.g., I felt dizzy; I felt my eyes were strained).

Social Attraction: The Social Attraction Questionnaire (SAQ; Nowak, 2004, adapted from McCroskey & McCain, 1974) was used to judge the likeability of the human-avatar. The SAQ includes 8 items rated on a 7-point Likert scale (*strongly agree – strongly disagree*) about how much the participant enjoyed their interaction with the avatar (e.g., I think they could be a friend of mine; I would like to have a friendly chat with them; I would like to meet them again).

Facial expression recognition: This task assessed the participant's ability to recognise emotions from static facial expressions presented within the CVE. Twenty-five greyscale face portraits were chosen from the series "Pictures of Facial Affect" (Ekman & Friesen, 1976), using the categories disgust, sad, angry, surprise and fear (cf. Wallace et al., 2008).that is, five faces from each category were shown.

Procedure

The VE was built on land purchased from Activeworlds Educational Universe (AWEDU©). Assistance in the design and features of the virtual world was gained from a young person with ASD on work experience at the Department of Psychiatry, University of Oxford. The virtual world consisted of several customised houses and galleries on a network of roads.

The participant and experimenter sat in different rooms at standard desktop computers and navigated the VE using keyboard controls. The child was told that the other avatar was controlled by an adult sitting in another room in the building. The gender, age and appearance of the avatars selected by the participant and the experimenter were similar. The virtual world was viewed from a first person perspective (i.e., through the eyes of the avatar) and participants could see the avatar of the person they were talking to onscreen. Participants and the experimenter wore microphone headsets so that they could talk, which was facilitated using Voice Over Internet Protocol (VOIP) software (see Figure 1).

Insert FIGURE 1 about here

1) Virtual World Training

Participants were taken to a room by one experimenter whilst another experimenter sat at a computer in a different room, ready to interact with the child. The experimenter accompanying the child guided them through a virtual tutorial prior to the experimental session, and from then on played a supervisory role; available to answer questions if required (which was rare). The avatar controlled by the experimenter in the other room (e-avatar) then greeted the participant on-screen with a wave and verbally through the headset, before engaging the participant in an exploration task (finding a present hidden in one of the houses) to allow the participant to practice the keyboard controls and acclimatise to the virtual world. All aspects of the interaction with the child throughout the session were semi-scripted to keep administration as consistent as possible across participants.

2) Positive/negative greeting - measuring social attraction

After training, participants met a friendly and an unfriendly character, both played by the e-avatar in interactions designed to measure impressions of positive versus negative greeting behaviour. Both characters were similar in appearance and were counterbalanced as to whether they would appear as a positive or negative greeter. The friendly character greeted the participant in an engaging manner (“My name is Paul. It’s very nice to meet you”) and initiated a conversation about school and hobbies (“What sort of things do you like doing outside of school?”). The unfriendly character initially did not approach the participant, did not initiate conversation and if the child made a social overture responded with short phrases with no vocal affect. Each of the two greeting scenarios lasted approximately 2 minutes and was counterbalanced in terms of which (positive or negative) the participant saw first. The participant answered the SAQ (Nowak, 2004), immediately after each greeting, which was administered by the experimenter sitting in the room with the participant. In total, this part of the procedure took about 10 minutes.

3) Facial Expression Recognition Task

Participants were first familiarised with the task by viewing three practice faces from the Ekman and Friesen (1976) series. Subsequently, the presentation order of faces in the virtual gallery was randomised for each participant. The names of the five emotions appeared in a frame alongside each face, and their order was also randomised for each face. Circles on the floor in front of each face indicated where the child’s avatar had to stand in order to view the picture, and a line under each face picture had to be aligned with the bottom of the computer screen to make sure each participant was at the same viewing distance. Participants were asked by the e-avatar to navigate themselves to the correct position in front of a picture and to tell them what emotion the face was showing from the five choices available; no other verbal prompts were given. Participants were then asked to navigate to the next picture and the procedure was repeated (Figure 2). Time taken varied between participants but took on average about 10 seconds per face or approximately 5 minutes to complete the full task.

Insert FIGURE 2 about here

4) Ratings of Presence

Once all tasks in the CVE were completed the experimenter in the room with the child administered the ITC-SoPI questionnaire (Lessiter et al., 2001). The experimenter helped the child if the meaning of any of the words was unclear, which was rare.

Results

Findings from individual measures

Sense of Presence: The four main sub-scales of the ITC-SoPI demonstrated good internal reliability: Spatial Presence $\alpha = 0.88$; Engagement $\alpha = 0.86$; Ecological Validity $\alpha = 0.67$; Negative Effects $\alpha = 0.73$. The mean scores for each group were calculated for these four main factors (see Table 2). Independent samples *t* tests showed that there were no statistical differences between the ASD and TD groups for any of the four factors: Spatial Presence (ASD, $M = 3.1$, $SD = .8$; TD, $M = 3.2$, $SD = 0.3$); Engagement (ASD, $M = 3.5$, $SD = .8$; TD, $M = 3.4$, $SD = .4$); Ecological Validity (ASD, $M = 2.8$, $SD = .9$; TD, $M = 2.6$, $SD = .5$); and Negative Effects (ASD, $M = 1.8$, $SD = .7$; TD, $M = 1.7$, $SD = .3$).

Social Attraction: Total social attraction scores from the SAQ (Nowak, 2004) were analysed for each group using a paired samples *t* test to compare responses to the positive and negative greeting scenarios. There was good internal reliability for these scales also: positive $\alpha = 0.86$ and negative $\alpha = 0.89$. The ASD group showed significant differences, $t(9) = 6.599$ $p < .001$, between their mean social attraction scores in response to the positive ($M = 40$, $SD = 7.6$) versus the negative greeting ($M = 33.1$, $SD = 7.3$). The TD group also showed significant differences, $t(9) = 6.854$ $p < .001$, in their mean social attraction scores in response to the positive ($M = 44.4$, $SD = 5.9$) versus the negative greeting ($M = 22.7$, $SD = 10.3$). The mean difference in social attraction scores between the two types of greeting was larger in the TD group compared to the ASD group, as shown by a significant group interaction with social attraction score (positive/negative), $F\,(1,19) = 19.704$ $p < .001$, (see Table 2).

Facial expression recognition: Percentage accuracy was calculated for recognition of each facial expression (disgust, sadness, anger, surprise and fear) and overall percentage accuracy was also calculated. Independent sample *t* tests showed that overall percentage accuracy was significantly lower, t (18) = 2.979 $p < .01$ d = 1.4, in the ASD group (M = 62.2, SD = 17.3) compared to the TD control group (M = 81.6, SD = 11.2). Further comparisons showed significant group differences for recognition of fear, t(18) = 3.162 $p < .01$ d=1.5 (ASD, M=28, SD = 28.5; TD, M= 66, SD= 25) and sadness t(18) = 2.240 $p < .05$ d=1.1 (ASD, M=64, SD = 26.3; TD, M= 86, SD = 16.5). Contrary to our prediction, no significant differences were found for recognition of disgust between the ASD (M=64, SD = 30) and TD (M=70, SD = 23.6) groups. No group differences were identified for recognising anger or surprise (see Table 2).

Insert TABLE 2 about here

Discussion

This paper is the first presentation of data on the self-reported feelings of presence of participants with ASD to social stimuli experienced in a Collaborative Virtual Environment (CVE) and a comparison of these with an age and IQ-matched group of TD participants. In line with previous research that tested levels of presence experienced by young people in an immersive VE room (Wallace et al., 2010), the current study found no differences between young people with ASD compared to TD controls in reported feelings of Spatial Presence, Engagement, Ecological Validity, and Negative Effects in a desktop VE. Thus, in a desktop VE young people with ASD felt spatially present inside the virtual world and engaged with an experience that they regarded as realistic. This finding suggests that, given an appropriate context and set of tasks, young people with ASD are willing to treat desktop CVEs as veridical media rather than simply as games with few links to reality (cf. Wallace et al., 2010; Parsons et al., 2004). Perhaps most importantly, young people with ASD, like their typically developing counterparts, reported low levels of negative effects overall (e.g., finding the content perceptually unpleasant). Taken together, these findings suggest that social responses and interpretation in the CVEs in this study are unlikely to be due to general experiences of the virtual medium (i.e., an artefact of 3-D technology).

Interestingly, both the ASD and TD groups found the unfriendly avatar in a greeting scenario significantly less socially attractive than the friendly avatar, based on their responses to the SAQ (Nowak, 2004). Wallace et al. (2010) reported that in a virtual playground scenario a misbehaving avatar (Danny) was not rated by young people with ASD as socially less attractive than a well-intentioned avatar. This was interpreted as reflecting a real world difficulty in understanding social intentions (cf. Baron-Cohen, 2000) that was reproduced in responses in the VE. The finding that the ASD group in the current study showed the typical difference in response to the two greeting scenarios may reflect the fact that the contrast between these greeting styles was deliberately distinctive, compared to the more nuanced conditions in Wallace et al. (2010).

Moreover, it is possible that the human-controlled avatar in the current study may have been interpreted as being more natural and therefore easier to interpret than the prerecorded animated avatar used in Wallace et al. (2010). This interpretation is supported by evidence, at least from TD groups, who judge human-controlled avatars and agent avatars differently (Blascovich et al., 2002), and anthropomorphic online characters to be more credible and attractive than less human-like characters (Nowak & Rauh, 2006). In addition, it could be that the dynamic nature of the interaction in the present study provided a more authentic opportunity in which participants with ASD could demonstrate their social competence (cf. Begeer, Rieffe, Meerum, Terwogt, & Stockmann, 2003; Cassidy, Mitchell, Chapman, & Ropar, 2015). Nevertheless, there was a significant group interaction with greeting scenario (positive versus negative) suggesting that sensitivity towards social initiations in this CVE task was reduced in the ASD compared to the TD group and therefore a CVE could be an ecologically valid context in which social difficulties are (at least partially) reproduced. The difference between the pattern of findings between the current study and Wallace et al. (2010) could be a cohort effect or related to the specific nature of the CVE task and so future work would need to conduct a direct comparison between tasks and different technology with the same groups in order to tease apart these factors.

Performance on the static Facial Expressions task was broadly in line with patterns of responding shown by participants taking part in traditional laboratory-based studies, that is,, using controlled computerised presentation but not using a CVE (Wallace et al., 2008). This confirms that the CVE set-up used in the present study provides an ecologically valid context in which difficulties with recognising facial affect can be replicated.

Limitations

As is typical in ASD technology research (Ploog, Scharf, Nelson, & Brooks, 2013), and indeed, psycho-educational ASD research more widely (Parsons et al., 2011), the sample size in this study is small and so the findings require further exploration and extension with larger groups. In addition, the interaction between the human avatar and the participant was brief, and within a single session, and so responses and interpretations may change over time, with repeated exposure. However, we suggest that a one-off session such as this provides an ecologically valid insight into initial impressions of CVE use, which may be akin to, for example, a single assessment session that could take place in a school or clinic. Of course, the extent to which the encounters in CVE can then inform us about likely responses to real world interactions remains an open question and more direct comparisons between real world and virtual responding need to be made (cf. Carter, Williams, Hodgins, & Lehman, 2014).

The use of static facial stimuli, rather than dynamic moving faces, may produce different findings in relation to impaired social judgements of facial expressions (Forgeot d'Arc et al., 2014) and so extending the CVE-based tasks to include dynamic stimuli will be a valuable next step. There were also no formal checks for procedural fidelity of the tasks and so there could be important, but subtle, differences in how the tasks were carried out that may have influenced responding. This is unlikely to be the case given that the researcher acting as the human avatar had a script to read from for the interactions, and the same researchers carried out all of the tasks with all of the participants, but is nevertheless something that would benefit from being demonstrated formally. Lastly, the participants included in this study had IQs within the normal range and so further work is required to extend this research to an exploration of how individuals with ASD and a learning disability use and respond to CVEs, not least because lower-functioning children seem to respond differently to more able children on VE-based tasks (Hopkins et al., 2011).

Implications for future research

Overall, ASD and TD participants reported similar levels of presence when interacting with human-controlled avatars in CVEs, but the groups responded differently on other measures exploring social responding and interpretations. This suggests that differential responding on socially oriented tasks is unlikely to be a technological artefact of CVEs and more likely to reflect real world difficulties in social cognition. Researchers and clinicians can take reassurance from the lack of negative effects, and similar levels of engagement and perceived ecological validity as TD participants, reported by participants with ASD with normal range IQs when using a CVE. Previous studies may have assumed this to be the case, but this is the first study to empirically demonstrate self-reported sense of presence by high-functioning participants with ASD in a CVE. For the future application of these technologies to be successful, researchers need to carefully consider their appropriateness, relevance, and usability in real world contexts. For example, Parsons et al. (2013) highlight the need to work in closer collaboration with educational practitioners when designing and implementing learning interventions. This is particularly true – and also especially challenging – for the technology field where education practitioners may be the ones having to learn new skills to work the technology or find ways to integrate hardware and software into their existing technology infrastructure. Consideration must also be given to the space technology will take up and how the cost will impact on school budgets (Grynszpan et al., 2014). For example, the desktop technology used in the present study is much more accessible and feasible for most schools than the technology used in Wallace et al. (2010). Thus, meaningful research–practitioner partnerships can help to prioritise needs and approaches within particular settings, for a particular cohort of students, and a clear plan can be developed from research to implementation.

Acknowledgments

We would like to thank the children who participated in this study and the teaching staff from the local school. We would also like to thank the young adult with autism who helped design the virtual world and who provided input into aspects of the project methods. Thanks also go to Heather Dilks and Sahil Suleman who contributed to this project as part of their undergraduate degree.

Author Note

This research was conducted without any external funding. The authors have no conflicts of interest in relation to this study.

References

Baron-Cohen, S. (2000). Theory of mind and autism: A review. *International review of research in mental retardation, 23*, 169–184. doi:10.1016/S0074-7750(00)80010-5

Begeer, S., Rieffe, C., Meerum Terwogt, M., & Stockmann, L. (2003). Theory of mind-based action in children from the autism spectrum. *Journal of Autism and Developmental Disorders*, 33, 479–487. doi:10.1023/A:1025875311062

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive virtual environment technology as a tool for social psychology. *Psychological Inquiry*, 13, 103–124. doi:10.1207/S15327965PLI1302_01

Bosseler, A., & Massaro, D. W. (2003). Development and evaluation of a computer-animated tutor for vocabulary and language learning in children with autism. *Journal of Autism and Developmental Disorders, 33*, 653–672. doi:10.1023/B:JADD.0000006002.82367.4f

Carter, E. J., Williams, D. L., Hodgins, J. K., & Lehman, J. F. (2014). Are children with autism more responsive to animated characters? A study of interactions with humans and human-controlled avatars. *Journal of Autism and Developmental Disorders, 44*, 2475–2485. doi:10.1007/s10803-014-2116-8

Cassidy, S., Mitchell, P., Chapman, P., & Ropar, D. (2015). Processing of spontaneous emotional responses in adolescents and adults with autism spectrum disorders: Effect of stimulus type. *Autism Research*. doi:10.1002/aur.1468

Charitos, D., Karadanos, E., Sereti, E., Triantafillou, S., Koukouvinou & Martakos, D. (2000). Employing virtual reality for aiding the organisation of autistic children behaviour in everyday tasks, *Proceedings of the 3rd International Conference on Disability, Virtual Reality and Associated Technologies (ICDVRAT)*, Sardinia, 23-25th September, 147–152.

Cheng, Y., & Ye, J. (2010). Exploring the social competence of students with autism spectrum conditions in a collaborative virtual learning environment–the pilot study. *Computers & Education*, *54*, 1068–1077. doi:10.1016/j.compedu.2009.10.011

Ekman, P., & Friesen, W.V. (1976). Pictures of facial affect. Palo Alto, CA,: Consulting Psychologists Press.

Fabri, M., & Moore, D.J. (2005). The use of emotionally expressive avatars in collaborative virtual environments, in Proceedings of Symposium on Empathic Interaction with Synthetic Characters, held at *Artificial Intelligence and Social Behaviour Convention 2005* (AISB 2005), University of Hertfordshire, April 2005.

Fabri, M., Elzouki, S., & Moore, D. (2007) Emotionally expressive avatars for chatting, learning and therapeutic intervention. *Proceedings of the 12th International Conference on Human-Computer Interaction (HCI International),* July 22-27, , Beijing, China, 4552, 275–285. doi:10.1007/978-3-540-73110-8_29

Forgeot d'Arc, B. F., Ramus, F., Lefebvre, A., Brottier, D., Zalla, T., Moukawane, S., … & Delorme, R. (2014). Atypical social judgment and sensitivity to perceptual cues in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 1–8. doi:10.1007/s10803–014–2208–5

Georgescu, A. L., Kuzmanovic, B., Roth, D., Bente, G., & Vogeley, K. (2014). The use of virtual characters to assess and train non-verbal communication in high-functioning autism. *Frontiers in Human Neuroscience*, *8* (article 807), 1–17 doi:10.3389/fnhum.2014.00807

Goodwin, M. S. (2008). Enhancing and accelerating the pace of autism research and treatment. *Focus on Autism and Other Developmental Disabilities*, *23*, 125–128. doi:10.1177/1088357608316678

Grynszpan, O., Nadel, J., Carbonell, N., Simonin, J., Constant, J., Le Bariller, F., …& Courgeon, M. (2009). A new virtual environment paradigm for high functioning autism intended to help attentional disengagement in a social context. *Proceedings of Virtual Rehabilitation (VR09),* University of Haifa, Israel, June 29th-July 2nd 2009, 51–58. doi:10.1109/ICVR.2009.5174205

Grynszpan, O., Weiss, P. L. T., Perez-Diaz, F., & Gal, E. (2014). Innovative technology–based interventions for autism spectrum disorders: A meta-analysis. *Autism*, 18, 346–61. doi:10.1177/1362361313476767

Hammick, J. K., & Lee, M. J. (2014). Do shy people feel less communication apprehension online? The effects of virtual reality on the relationship between personality characteristics and communication outcomes. *Computers in Human Behavior*, *33*, 302–310. doi:10.1016/j.chb.2013.01.046

Hopkins, I. M., Gower, M. W., Perez, T. A., Smith, D. S., Amthor, F. R., Wimsatt, F. C., & Biasini, F. J. (2011). Avatar assistant: improving social skills in students with an ASD through a computer–based intervention. *Journal of Autism and Developmental Disorders*, *41*, 1543–1555. doi:10.1007/s10803-011-1179-z

Jacobson, D. (2001). Presence revisited: Imagination, competence, and activity in text-based virtual worlds. *CyberPsychology & Behavior*, *4*, 653–673. doi:10.1089/109493101753376605

Josman, N., Milika, Ben-Chaim. H., Friedrich, S. & Weiss, P. L. (2008) Effectiveness of virtual reality for teaching street-crossing skills to children and adolescents with autism. *International Journal on Disability and Human Development,* 7, 49–56. doi:10.1515/IJDHD.2008.7.1.49

Kandalaft, M. R., Didehbani, N., Krawczyk, D. C., Allen, T. T., & Chapman, S. B. (2013). Virtual reality social cognition training for young adults with high-functioning autism. *Journal of Autism and Developmental Disorders*, *43*, 34–44. doi:10.1007/s10803-012-1544-6

Ke, F., & Im, T. (2013). Virtual-reality-based social interaction training for children with high-functioning autism. *The Journal of Educational Research*, *106*, 441–461. doi:10.1080/00220671.2013.832999

Knight, V., McKissick, B. R., & Saunders, A. (2013). A review of technology-based interventions to teach academic skills to students with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *43*, 2628–2648. doi:10.1007/s10803-013-1814-y

Leekam, S. R., Nieto, C., Libby, S. J., Wing, L., & Gould, J. (2007). Describing the sensory abnormalities of children and adults with autism. *Journal of Autism and Developmental Disorders*, *37*, 894–910. doi:10.1007/s10803-006-0218-7

Lessiter, J., Freeman, J., Keogh, E. & Davidoff, J. (2001). A cross-media presence questionnaire: the ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments, 10*, 282–297. doi:10.1162/105474601300343612

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview—Revised. *Journal of Autism and Developmental Disorders, 24*, 659–685. doi:10.1007/BF02172145

Massaro, D. W. & Bosseler, A. (2006). Read my lips: The importance of the face in a computer-animated tutor for vocabulary learning by children with autism. *Autism*, 10, 495–510. doi:10.1177/1362361306066599

McCroskey, J.C. & McCain, T.A (1974). The measurement of interpersonal attraction. *Speech Monographs*, 41, 261–266. doi:10.1080/03637757409375845

Mikropoulos, T. A. (2006). Presence: a unique characteristic in educational virtual environments. *Virtual Reality*, *10*, 197–206. doi:10.1007/s10055-006-0039-1

Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education*, *56*, 769–780. doi:10.1016/j.compedu.2010.10.020

Moore, D., Cheng, Y., McGrath, P. & Powell, N. (2005). Collaborative virtual environment technology for people with autism. *Focus on Autism and Other Developmental Disabilities*, 20, 231–243. doi:10.1177/10883576050200040501

Mottron, L., Dawson, M., Soulieres, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders, 36*, 27–43. doi:10.1007/s10803-005-0040-7

Nowak, K., & Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence*, *12*, 481–494 . doi:10.1162/105474603322761289

Nowak, K. L. (2004). The influence of anthropomorphism and agency on social judgment in virtual environments. *Journal of Computer-Mediated Communication*, 9. doi:10.1111/j.1083-6101.2004.tb00284.x

Nowak, K. L., & Rauh, C. (2006). The influence of the avatar on online perceptions of anthropomorphism, androgyny, credibility, homophily, and attraction. *Journal of Computer-Mediated Communication*, *11*, 153–178. doi:10.1111/j.1083-6101.2006.tb00308.x

Parsons, S., Charman, T., Faulkner, R., Ragan, J., Wallace, S. & Wittemeyer, K. (2013) Bridging the research and practice gap in autism: the importance of creating research partnerships with schools. *Autism*, 17, 268–280. doi:10.1177/1362361312472068

Parsons, S. & Cobb, S. (2011) State-of-the-art of virtual reality technologies for children on the autism spectrum. *European Journal of Special Needs Education*, 26, 355–366. doi:10.1080/08856257.2011.593831

Parsons, S., Guldberg, K., MacLeod, A., Jones, G., Prunty, A., & Balfe, T. (2011). International review of the evidence on best practice in educational provision for children on the autism spectrum. *European Journal of Special Needs Education*, *26*, 47–63. doi:10.1080/08856257.2011.543532

Parsons, S., Leonard, A. & Mitchell, P. (2006). Virtual environments for social skills training: comments from two adolescents with autistic spectrum disorder. *Computers & Education*, 47, 186–206. doi:10.1016/j.compedu.2004.10.003

Parsons, S. & Mitchell, P. (2002). The potential of virtual reality in social skills training for people with autistic spectrum disorders. *Journal of Intellectual Disability Research*, 46, 430–443. doi:10.1046/j.1365-2788.2002.00425.x

Parsons, S., Mitchell, P., & Leonard, A. (2004). The use and understanding of virtual environments by adolescents with autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, 34, 449–466. doi:10.1023/B:JADD.0000037421.98517.8d

Parsons, S., Mitchell, P. & Leonard, A. (2005). Do adolescents with autistic spectrum disorders adhere to social conventions in virtual environments? *Autism*, 9, 95 -117. doi:10.1177/1362361305049032

Ploog, B. O., Scharf, A., Nelson, D., & Brooks, P. J. (2013). Use of computer-assisted technologies (CAT) to enhance social, communicative, and language development in children with autism spectrum disorders. *Journal of Autism and Developmental disorders*, *43*, 301–322. doi:10.1007/s10803-012-1571-3

Riby, D., & Hancock, P. J. (2009). Looking at movies and cartoons: eye-tracking evidence from Williams syndrome and autism. *Journal of Intellectual Disability Research, 53*, 169–181. doi:10.1111/j.1365-2788.2008.01142.x

Rosset, D. B., Rondan, C., Da Fonseca, D., Santos, A., Assouline, B., & Deruelle, C. (2008). Typical emotion processing for cartoon but not for real faces in children with autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 38*, 919–925. doi.org/10.1007/s10803-007-0465-2

Schmidt, M., Laffey, J. M., Schmidt, C. T., Wang, X., & Stichter, J. (2012). Developing methods for understanding social behavior in a 3D virtual learning environment. *Computers in Human Behavior*, 28, 405–413. doi:10.1016/j.chb.2011.10.011

Schroeder, R. (Ed.). (2002). *The social life of avatars: Presence and interaction in shared virtual environments.* London: Springer-Verlag. doi:10.1007/978-1-4471-0277-9

Schwartz, C., Bente, G., Gawronski, A., Schilbach, L., & Vogeley, K. (2010). Responses to nonverbal behaviour of dynamic virtual characters in high-functioning autism. *Journal of Autism and Developmental Disorders*, *40*, 100–111. doi:10.1007/s10803-009-0843-z

Slater, M. (2004). How colorful was your day? Why questionnaires cannot assess presence in virtual environments. *Presence*, *13*, 484–493. doi:10.1162/1054746041944849

Stendal, K., Balandin, S., & Molka-Danielsen, J. (2011). Virtual worlds: A new opportunity for people with lifelong disability?. *Journal of Intellectual and Developmental Disability*, *36*, 80-83. doi:10.3109/13668250.2011.526597

Stichter, J. P., Laffey, J., Galyen, K., & Herzog, M. (2014). iSocial: Delivering the social competence intervention for adolescents (SCI-A) in a 3D virtual learning environment for youth with high functioning autism. *Journal of Autism and Developmental Disorders*, *44*, 417–430. doi:10.1007/s10803-013-1881-0

Strickland, D., Marcus, L, M., Mesibov, G. B. & Hogan, K. (1996). Brief report: Two case studies using virtual reality as a learning tool for autistic children. *Journal of Autism and Developmental Disorders*, 26 , 651–659. doi:10.1007/BF02172354

Strickland, D., McAllister, D., Coles, C.D., & Osborne, S. (2007). An evolution of virtual reality training designs for children with autism and fetal alcohol spectrum disorders. *Topics in Language Disorders*, 27, 226–241. doi:10.1097/01.TLD.0000285357.95426.72

Trepagnier, C.G. (1999). Virtual environments for the investigation and rehabilitation of cognitive and perceptual impairments. *NeuroRehabilitation*, 12, 63–72.

Wallace, S., Coleman, M., & Bailey, A. (2008). An investigation of basic facial expression recognition in autism spectrum disorders. *Cognition & Emotion, 22,* 1353–1380. doi:10.1080/02699930701782153

Wallace, S., Parsons, S., Westbury, A., White, K., White, K., & Bailey, A. (2010). Sense of presence and atypical social judgments in immersive virtual reality: Responses of adolescents with Autistic Spectrum Disorders. *Autism,* 14, 199–213. doi:10.1177/1362361310363283Waller, D. (2000). Individual differences in spatial learning from computer-simulated environments. *Journal of Experimental Psychology: Applied*, *6*, 307. doi:10.1037/1076-898X.6.4.307

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence–Revised*. The Psychological Corporation.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization.

Yee, N., Bailenson, J. N., Urbanek, M., Chang, F., & Merget, D. (2007). The unbearable likeness of being digital: The persistence of nonverbal social norms in online virtual environments. *CyberPsychology & Behavior*, *10*, 115–121. doi:10.1089/cpb.2006.9984

Tables

Table 1. Background characteristics and matching details of the ASD and typically developing (TD) groups according to age and IQ (VIQ, PIQ & FIQ).

	ASD (*n* = 10)			**Typically Developing (*n* = 10)**			**Comparison between ASD and TD**
	M	*Range*	*SD*	*M*	*Range*	*SD*	
Age	14.8	12–16 yrs	1.1	15.4	14–16 yrs	0.8	*ns*
Verbal IQ	99	78–120	13.5	94.5	87–117	10.3	*ns*
Performance IQ	102	82–127	16.5	100.5	86–114	8.0	*ns*
Full IQ	100	79–124	15.7	97.4	86–113	8.7	*ns*

Table 2: Mean (M) scores and standard deviations (SD) by group for factors spatial presence, engagement, ecological validity and negative effects from the ITC-SoPI (Lessiter et al., 2001), and positive and negative greeting scores from the Social Attraction Questionnaire (Nowak 2004). Facial expression scores are mean percentage accuracy for each facial expression tested.

	ASD (n=10)		**TD** (n=10)		**Comparison between ASD and TD**
	M	*SD*	*M*	*SD*	
ITC-SoPI					
Spatial Presence	3.1	0.8	3.2	0.3	*ns*
Engagement	3.5	0.8	3.4	0.4	*ns*
Ecological Validity	2.8	0.9	2.6	0.5	*ns*
Negative Effects	1.8	0.7	1.7	0.3	*ns*
Social attraction scores					
Positive greeting	40	7.6	44.4	5.9	*ns*
Negative greeting	33.1	7.3	22.7	10.3	*ns*
Group Interaction [positive*negative]					$p < .001$
Facial expression recognition					
Disgust	64	30.0	70	23.6	*ns*
Sadness	64	26.3	86	16.5	$p < .05$
Anger	74	19.0	88	14.0	*ns*
Surprise	82	25.7	98	6.3	*ns*
Fear	28	28.6	66	25.0	$p < .01$
Overall score	62.2	17.3	81.6	11.2	$p < .01$

Figure captions

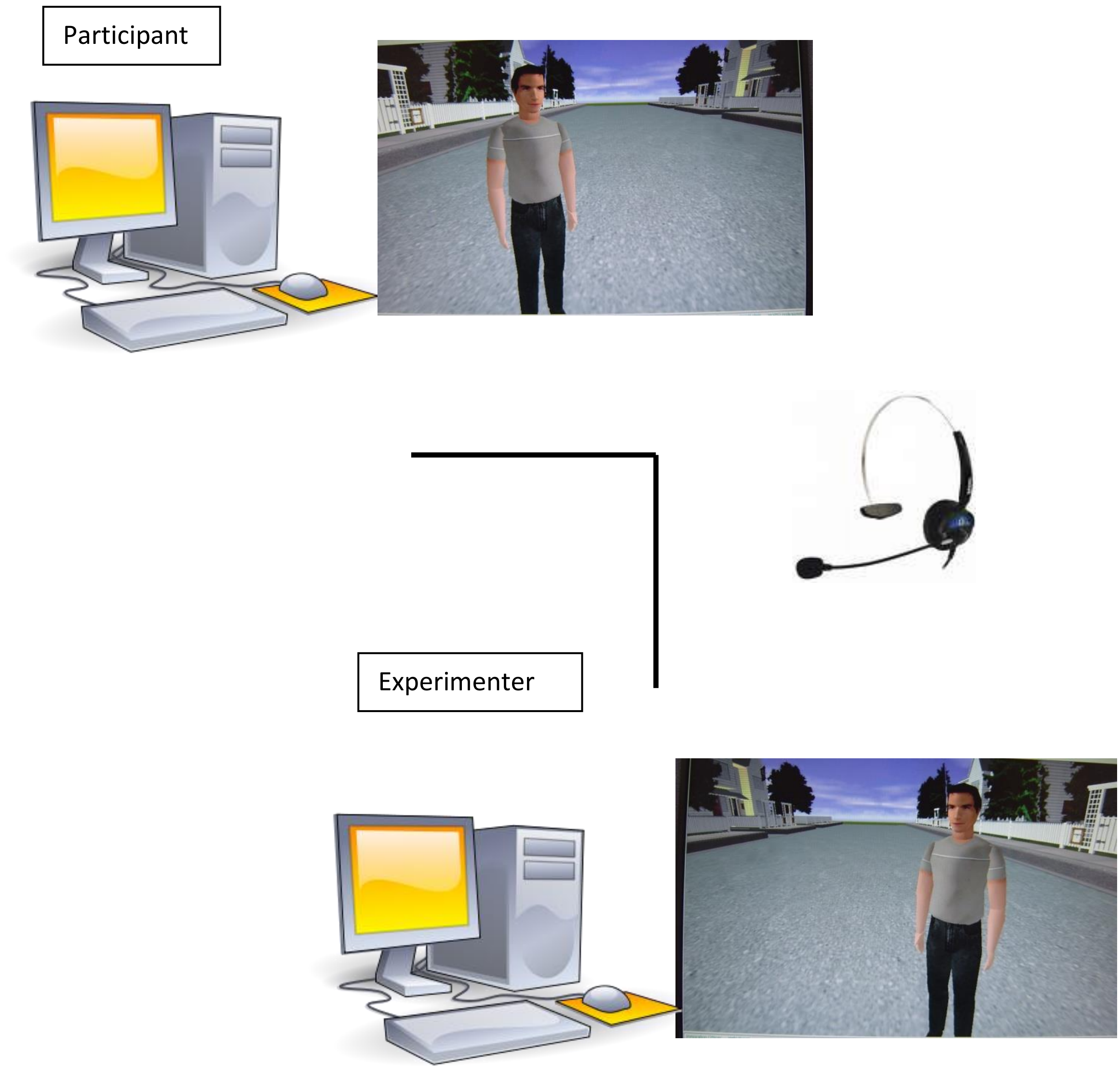

Figure 1. Participants and experimenter sat in different rooms and were connected via desktop computers into the virtual world where they each appeared as avatars. The virtual world was viewed in the first person perspective and the participant and experimenter communicated via voice over internet protocol (VOIP) using microphone headsets.

Images from the Activeworlds Educational Universe are used with the permission of www.activeworlds.com (where more recent examples of VEs can be accessed)

Figure 2. The virtual gallery where participants were asked to navigate themselves in front of a picture frame and verbally identify the facial expression using the labels that appeared next to the frame