Computers & Education 117 (2018) 1–15

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Enhance emotional and social adaptation skills for children with autism spectrum disorder: A virtual reality enabled approach

Horace H.S. Ip [a], Simpson W.L. Wong [b], Dorothy F.Y. Chan [c], Julia Byrne [a], Chen Li [a, *], Vanessa S.N. Yuan [a], Kate S.Y. Lau [a], Joe Y.W. Wong [a]

[a] Centre for Innovative Applications of Internet and Multimedia Technologies, City University of Hong Kong, Kowloon Tong, Hong Kong
[b] Department of Education Studies, Hong Kong Baptist University, Kowloon Tong, Hong Kong
[c] Department of Paediatrics, Faculty of Medicine, The Chinese University of Hong Kong, Shatin, Hong Kong

ARTICLE INFO

Article history:
Received 30 March 2017
Received in revised form 17 September 2017
Accepted 23 September 2017
Available online 27 September 2017

Keywords:
Autism spectrum disorders
Virtual reality
Emotional skills
Social adaptation
Situated learning

ABSTRACT

Deficits in social-emotional reciprocity, one of the diagnostic criteria of Autism Spectrum Disorder (ASD), greatly hinders children with ASD from responding appropriately and adapting themselves in various social situations. Although evidences have shown that virtual reality environment is a promising tool for emotional and social adaptation skills training on ASD population, there is a lack of large-scale trials with intensive evaluations to support such findings. This paper presents a virtual reality enabled program for enhancing emotional and social adaptation skills for children with ASD. Six unique learning scenarios, of which one focuses on emotion control and relaxation strategies, four that simulate various social situations, and one that facilitates consolidation and generalization, are designed and developed with corresponding psychoeducation procedures and protocols. The learning scenarios are presented to the children via a 4-side immersive virtual reality environment (a.k.a., half-CAVE) with non-intrusive motion tracking. A total number of 94 children between the ages of 6–12 with clinical diagnosis of ASD participated in the 28-session program that lasted for 14 weeks. By comparing pre- and post-assessments, results reported in this paper show significant improvements in the project's primary measures on children's emotion expression and regulation and social-emotional reciprocity but not on other secondary measures.

1. Introduction

Persistent deficits in social-emotional reciprocity is the important diagnostic criterion for Autism Spectrum Disorder (ASD) (American Psychiatric Association, 2013). For school-aged children with ASD, deficits in emotional and social adaptation skills greatly hinder them from adapting themselves to the school setting. In Hong Kong, a school-aged child with ASD and an IQ above the cut-off range is required to attend mainstream school in the inclusive education setting. With the belief that all children have equal access to quality education, one of the key objectives of inclusive education is to help children with ASD to

* Corresponding author.
E-mail address: richard_lichen@hotmail.com (C. Li).

https://doi.org/10.1016/j.compedu.2017.09.010
0360-1315/

better integrate into the society at young ages. However, without proper guidance and resources, this not only brings challenges to the children themselves, but also to their peers and teachers who may not have enough professional supports.

Virtual Reality (VR) has recently been studied and discovered as a promising tool to address the psychoeducational needs of children with ASD in various settings. These studies cover a great range of training and interventions, including emotional skills training (e.g. Bekele et al., 2014) and social adaptation training (e.g. Kandalaft, Didehbani, Krawczyk, Allen, & Chapman, 2013; Smith et al., 2014). As reported in these studies, the VR environment can provide safe and controlled yet highly interactive and realistic virtual scenarios, which can well facilitate the psychoeducational needs of people with ASD. More specifically, the computer-generated graphics attract the attention and interest of people with ASD because many of them primarily rely on visual thinking (Jacobson, 1978). Moreover, because social situations can be repeatedly reproduced in the virtual reality environment, they could be introduced adaptively for knowledge and skill generalization (Parsons & Mitchell, 2002; Strickland, Marcus, Mesibov, & Hogan, 1996).

Since 1990s, many empirical research studies have investigated the feasibility and effectiveness of using virtual reality for training and intervention on the ASD population (Parsons & Cobb, 2011). The reported results are promising. However, certain issues and limitations need to be further addressed. Above all, most of these research studies only focus on interventions when subjects are being exposed to VR while after-intervention protocols and methods, if they do exist, were not well reported. We believe after-intervention activities are as important as VR-enabled intervention itself. Besides, the VR learning contents reported in previous research studies are too specific in terms of simulated scenarios and situations. To fully unleash the potential of virtual reality enabled ASD training and intervention and better facilitate generalization, simulated scenarios need to authentic, rich, and flexible; this is especially critical for social adaptation trainings (Parsons & Cobb, 2011). Finally, the numbers of subjects in previous studies are relatively small, which limits the authors' ability to draw meaningful conclusions from quantitative statistics based on well-designed experiments (McCann et al., 2014). The work reported in this paper focuses not only on the six uniquely designed virtual reality scenarios which cover most of the real scenes for school-aged children during school days, but also the protocols and methods to support the children after intervention. A total number of 94 children with clinical diagnoses of ASD participated in this empirical study. This is, as we know it, the largest empirical study to date on using immersive VR technologies on psychoeducation for children with ASD, in terms of number of participants.

In this paper, a VR-enabled training program for enhancing emotional and social adaptation skills for children with ASD is presented with evaluation results. We hypothesize that children with ASD would improve on their emotional and social adaptation skills after the training. More precisely, we will be looking at the children's skills and ability on emotion recognition, emotion expression and regulation, social interaction and adaptive skills. Our primary outcomes focus on emotion expression and regulation, and social interaction, while our secondary outcomes include emotion recognition and adaptive skills.

This paper is organized as follow: section 2 provides a literature review relevant to this study, namely the social-emotional reciprocity in children with ASD and traditional interventions, immersive virtual reality technologies, and previous studies that adopt virtual reality technologies, especially immersive virtual reality technologies for ASD training and intervention. In section 3, details regarding the immersive virtual reality system are described, including hardware environment, software environment, and content design, followed by method and training procedures presented in section 4. Section 5 presents the assessment data, which are discussed in section 6. We conclude our findings and propose future works in section 7.

2. Related works

2.1. Social-emotional reciprocity in children with ASD and traditional interventions

A diagnosis of ASD composes of persistent impairment and deficits in social communication and patterns of repetitive behaviors (American Psychiatric Association, 2013). Children with ASD varies differently in their behaviors but many of them have difficulties in social skills and emotion regulation. Some social communication impairments include lack of social reciprocity and emotion reciprocity, poor eye contact and difficulties in starting and maintaining conversation. A lot of these children also have sensory-perceptual abnormalities that may add on to these difficulties. These impairments contribute to the everyday struggles of these children with ASD and their family. Therefore, it is important to understand these underlying difficulties to develop and deliver appropriate interventions to help them.

Social interaction is a reciprocal process in which all individuals participate in an active process of initiating and responding to social stimuli with their peers (Shores, 1987). It is an essential skill that starts to emerge at young age and is highly required in many daily life settings, such as in class, on the playground, or during mealtime. Children with ASD participate in less social interactions than their typically developed peers, while their interactions are also often poor in quality and considered inappropriate according to social norms (Sigman & Ruskin, 1999).

Research has found that between 30% and 100% of people with ASD have sensory-perceptual abnormalities (Dawson & Watling, 2000). These abnormalities include sensory distortion, sensory overload, hyposensitivity and hypersensitivity, inappropriate responses to multiple sensory stimuli, etc. (Watling, Deitz, & White, 2001). Specifically, people with autism report having sensory overload and difficulties with stimulation from more than one channel at the same time (O'neill & Jones, 1997). A study has also found that sensory processing is a function of severity in ASD, while there are relationships between sensory processing levels and social responsiveness, a causal relationship has yet to be developed (Hilton, Graver, & LaVesser, 2007).

In order to help these individuals from a young age, a number of psychosocial interventions have been developed, such as Treatment and Education of Autistic and Related Communication Handicapped Children (TEACCH), Applied Behavior Analysis (ABA), social skills training, etc. (Foxx, 2008; Mesibov, Shea, & Schopler, 2005). For instance, social stories that are composed of different scenarios were designed to help individuals with ASD to better understand themselves and others (Gray & Garand, 1993). With improved understanding, children may respond more appropriately in actual social settings.

Furthermore, Computer-Based Intervention (CBI), is also widely used in enhancing social communication of children with ASD (Higgins & Boone, 1996). There is also increasing evidence on virtual reality (VR) as a new approach to social skills training. It accommodates the strengths of traditional social skills training while allowing for more possibilities in a safe, controllable environment (Parsons & Mitchell, 2002). The cognitive theory of multimedia learning states that people learn deeper and better from words and pictures together than from words solely, making VR a promising medium for intervention delivery (Mayer, 2002). It is important to ensure that both auditory and visual information are coherently presented while neither of them would cause sensory overload (Mayer, 2002). The close matchup between actions and dialogues in the virtual reality environment provides an effective learning gateway.

2.2. Immersive virtual reality technologies

Steuer (1992) defines VR from the perspective of user experience by emphasizing *the sense of presence* while being exposed to VR. An immersive virtual reality environment should be able to let its users feel they are *in* and *a part of* the computer-generated virtual scenarios. This requires the VR environment to provide not only vivid multi-sensory stimulations but also high interactivity to its users, both of which are supposed to be challenging given the currently available and accessible technologies. Currently, immersive virtual reality experience is dominated by visual and audial stimulations and the interactivity is becoming more and more intuitive, thanks to the new generation of sensor technologies.

Head-mounted display (HMD) is a type of virtual reality display which integrates two displays and corresponding lenses for the left eye and the right eye respectively in order to enable stereoscopic viewing. Early generations of HMDs are bulky and heavy, and could only display low resolution images. Latency is another issue that needs to be considered when employing HMDs because latency is believed to be the major cause for visual system discomfort (Merhi, Faugloire, Flanagan, & Stoffregen, 2007; Sharples, Cobb, Moody, & Wilson, 2008). Although the new generation of HMDs, such as HTC VIVE and Oculus Rift, can greatly improve the user experience by reducing weight, increasing display resolution, and decreasing latency, the uniqueness of HMD still needs to be taken into account when employing HMDs in the context of ASD training and intervention; the HMD visually isolates people with ASD completely from the outsiders, including the training facilitator, which may negatively affect the effectiveness and even cause safety concerns. Projection-based immersive virtual reality systems have their advantages over HMDs particularly for the concerns mentioned above given the context of employing virtual reality technologies for ASD training and intervention. A projection-based immersive virtual reality system usually blends two or more stereoscopic screens together to provide the immersive experience. Cave Automatic Virtual Environment (CAVE) (Cruz-Neira, Sandin, DeFanti, Kenyon, & Hart, 1992) is a projection-based immersive virtual reality system of which projection screens are arranged in a cubic configuration. The images displayed on each screen are generated by the rendering workstation(s) based on the position and orientation of the viewer's head, so that the viewing perspective can be adjusted automatically in real-time. Cruz-Neira, Sandin, and DeFanti (1993) described the mathematical details regarding the computer graphics rendering in a CAVE setting. DeFanti et al. (2009) provided a hexagonal variant to the original box-shaped CAVE, which claimed to be able to provide better virtual reality experience to users in the hexagonal prism space.

Sensors, particularly motion sensors, are the key to enable intuitive interactions in virtual reality environments. Ip, Hay, and Tang (2002) applied computer vision algorithm to the raw images captured by two infrared cameras in order to enable non-intrusive motion tracking in a projection-based immersive VR environment for creative arts. The enhanced version of this technology was later applied for interactive education program (Ron, Cheng, Ip, & Joseph, 2011). More recent inventions, such as Kinect (Shotton et al., 2013) and LeapMotion focusing on whole body gesture recognition and hand gesture cognition respectively, are also based on infrared camera enabled computer vision solution but packed in a much smaller and more accessible form factor. These motion sensors have been used in VR systems for ASD training and intervention (Cai et al., 2013; Lorenzo, Pomares, & Lledó, 2013).

2.3. Virtual reality enabled training approaches for ASD children

Virtual reality and its enabling technology is believed to be an effective tool to facilitate trainings for children with ASD in various domains, ranging from generic social skills (e.g. Kandalaft et al., 2013) to more specific tasks such as job interviewing skills (e.g. Smith et al., 2014), coping with specific anxieties (e.g. Rothbaum et al., 2006), safety awareness and education (e.g. Matsentidou & Poullis, 2014), etc. By employing VR technologies, the highly interactive virtual environment which mimics real scenes can provide the sense of presence to the children while keeping them away from potential danger or unnecessary embarrassment. Moreover, the scenes can be controlled and manipulated by the professionals to provide variants during repetitive training sessions in order to help the children to generalize knowledge and skills to new situations in both virtual and real contexts. Here, we review relevant research studies on VR-enabled psychoeducational approaches for children with ASD.

The early adoption of virtual reality as a learning tool for children with ASD can be traced back to 1990s. In 1996, a study attempted to investigate the potential of using VR as a learning tool for children with ASD (Strickland et al., 1996). Two ASD children, a 7.5-year-old girl and a 9-year-old boy, participated in this study with the facilitations from their parents. As reported by the authors, both children accepted the relatively bulky and heavy head-mounted display (HMD) and were able to identify familiar objects and their quantities in a real-time computer-generated street scene. This study shows that children with ASD could accept and respond to the virtual scenarios in a meaningful way, which is the very fundamental precondition of adopting VR for educational purposes. However, due to the limitations of VR technologies at the time, the computer-generated scenarios tended to be primitive in terms of visual details, and the child-VR interactions were not very intuitive. Compared to this study, other researchers later reported similar findings by using café and bus virtual environments with more visual details (Parsons, Leonard, & Mitchell, 2006; Strickland et al., 1996). In summary, even in the very early days of VR when the enabling technologies could barely provide real-time computer generated virtual contents and very limited interactions, there were evidences showing that children with ASD could accept and respond to computer-generated virtual environments.

Recent studies adopt more advanced VR technologies to increase both the vividness and interactivity of the virtual environments, and to enable more intuitive user-VR interactions. One of those studies combined immersive display technology with non-intrusive sensor technologies for creating an immersive VR system in order to support the learning of students with Asperger syndrome (Lorenzo et al., 2013). The virtual learning environment was presented via a two-side L-shape stereoscopic projection system with automatic perspective correction (a.k.a., semi-CAVE) achieved through its positioning system. Other sensors, including Microsoft Kinect and a set of additional cameras, were also part of the virtual reality installation. The Kinect sensor allowed students to interact with the virtual scene in an intuitive way and the additional camera system could recognize and record students' facial expressions during learning. Ten primary school students and ten secondary school students, all with Asperger syndrome, participated in this study. Results showed that after the intervention, participants' social competences and executive functions had improved. Matsentidou and Poullis (2014) employed an even more immersive VR installation, a four-side rear projection screens with automatic perspective correction (a.k.a., half-CAVE). A virtual scenario that mimicked the real-life situation of crossing the road called "Lost in City" was designed and developed to train children with ASD how to react appropriately when facing unsafe situations. User-VR interactions were mainly achieved via Xbox controller - a common gamepad with joysticks and buttons designed to ease the navigations and interactions in games. Twelve children between 9 and 10 years old participated in the evaluation of the system. The reported results were encouraging. However, they were mostly based on qualitative observations rather than quantitative analysis. The 320-degree cylindrical installation blending five stereoscopic projectors employed in (Cai et al., 2013) is another way of achieving immersive VR experience other than cubic CAVE installations. The authors used Microsoft Kinect sensor to recognize hand gestures for dolphin training, and successfully transferred the dolphin-assisted therapy approach from aquariums to the virtual environment. Fifteen children with ages ranging from 6 to 17 years old were randomly selected from a pool of 32 children with ASD. They all have gone through real dolphin-assistant therapy between 2007 and 2011. Most of them showed a certain degree of interest in the virtual dolphins in the virtual scenario, and were willing to interact with the virtual dolphins following the video tutorials on hand gestures during dolphin training.

Immersive VR environment, as a new medium to facilitate ASD training and intervention, is now starting to exhibit its potentials, partly due to its enabling technologies are becoming more and more accessible. However, like any new approaches in their early stages and particularly with ASD population, there is still a lack of systematic studies conducted on a relatively large number of subjects. The study reported in this paper adopt VR-enabled approaches to training for and evaluate it on a much larger population, in this case, 94 children with clinical diagnosis of ASD.

3. Immersive virtual reality system

The main objective of creating a VR-enabled system for ASD training is to facilitate the trainers with simulated virtual scenarios in which situated learning could be applied. The hardware and software environment, together with the VR contents, are designed and developed with three fundamental principles in mind.

(1) The VR scenarios should contain enough visual details and visual cues, and should be presented to the children in an immersive VR environment with great fidelity;
(2) The VR scenarios should cover real life situations experienced by school-aged children in Hong Kong in their daily life;
(3) While giving the children sufficient freedom to interact with objects and avatars in the virtual scenarios, the trainers should still be able to lead and maintain the psychoeducation procedures and intervene when it is necessary.

Considering these principles, a Cave Automatic Virtual Environment (CAVE) (Cruz-Neira et al., 1992) is chosen as the enabling technology to facilitate the designated program. In this section, the details regarding hardware environment, software environment, and VR content design and development are described.

3.1. Hardware environment

Three rear projection screens and one front projection screen with four corresponding stereoscopic projectors are installed to assemble the four-side CAVE VR system (a.k.a., half-CAVE). Each of the four screens is a 2.75 m × 2.75 m square, forming a 2.75 m (W) x 2.75 m (D) x 2.75 m (H) cubic VR environment (see Fig. 1). Projection mirrors are installed to shorten the needs of horizontal rear-projection distance, hence much less floor space is occupied by the half-CAVE. The four stereoscopic projectors, each running at 1024 × 768 resolution and 120 Hz refresh rate, are driven by the rendering workstation equipped with two Intel Xeon E5-2687W v3 processors, two NVIDIA Quadro M6000 12GB graphics cards, one NVIDIA Sync card for stereoscopic signal synchronization, and a total number of 128GB of system RAM. In order to achieve automatic perspective correction in the half-CAVE, the motion tracking system, which consists of 20 sets of OptiTrack Flex 13 IR-based optical motion capture (MOCAP) cameras connecting to a dedicated workstation via USB connections, is also installed to track the target attached to the stereoscopic viewing goggles. The two workstations, namely the rendering workstation and the MOCAP workstation, are connected via the 1-Gbps local area network (LAN) for data exchanging.

3.2. Software environment

To ease the software development process, MiddleVR for Unity3D is used to handle challenges introduced by adopting CAVE VR system, including screen configuration, automatic perspective correction, projection edge blending, etc. The middleware is configured to fully reflect the physical installation of the hardware environment. This allows the viewing perspective could be adjusted automatically in real-time by modifying the projection matrix of each of the four virtual cameras in Unity3D, given the data being fetched simultaneously from the motion tracking system via VRPN protocol (Taylor II et al., 2001). The virtual reality contents are developed and tested in Unity3D 5.1 using C# as the programming language.

3.3. Content design

The VR content, in the form of learning scenarios, is designed to provide a controlled, safe and authentic environment for children with ASD to develop and practice their emotional and social adaptation skills facilitated by the trainer. To achieve the designated objectives, six unique scenarios have been designed and developed, of which one focuses on emotion control and relaxation strategies, four that simulate various social situations, and one that facilitates consolidation and generalization. Specifically, in the emotion control and relaxation scenario, children will be able to experience the four seasons and interact with virtual objects in the VR environment. This scenario is designed to help children getting used to the immersive virtual reality environment during the first two training sessions and provide a peaceful environment for them to calm down during emotional outbursts in the following sessions. Besides that, scenario 1 recreates the home scene in which children can practice morning routines on preparing to go to school. The scenario begins in the bedroom when the alarm clock was set off. The child is encouraged to navigate to the alarm clock and stop it by tapping. A checklist will pop up to indicate task completion and to direct the child to the next step, which is using the washroom. A social situation, such as the washroom is being occupied, can be introduced by the trainer. In that case, the child needs to wait patiently at the washroom door until an avatar opens the bathroom door and leave. Next, the child is guided to take care of his/her hygiene. The child is asked to

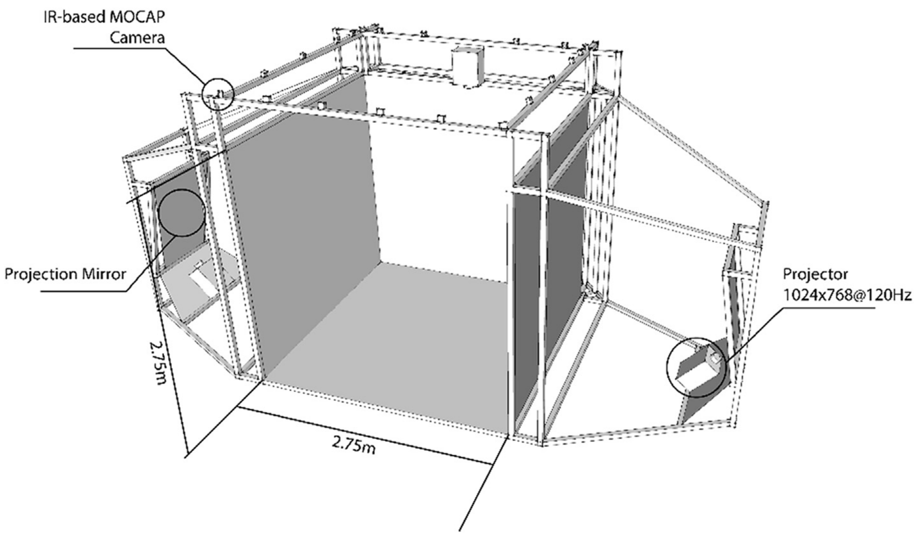


Fig. 1. Illustration of the half-CAVE virtual reality environment.

practice hand washing and teeth brushing, before he/she gets dressed and has breakfast. After taking the schoolbag at the doorway and changing his/her shoes, the child needs to take the lift to the lobby. Here, a social situation, such as the lift carriage is full, can be introduced under the supervision of the trainer. The child is encouraged to wait patiently for a lift carriage that is not full. The child will then take the lift to the lobby, greet the security staff at the main door, and wait for the school bus. This is where scenario 1 ends. Following scenario 1, scenario 2 simulates the scene of taking school bus to the school and having a general education class in the classroom. Similar to scenario 1, scenario 2 also consists a series of routines that children can practice and experience during school days in Hong Kong's inclusive education setting, including greet the teacher, hand in homework, follow the teacher's instruction and join learning activities, etc. The emotion recognition practice is embedded as a game-based learning activity during the class. Scenario 3 recreates the school library, while it is particularly challenging for children with ASD to stay socially appropriate in such settings. Children will be asked to follow the library rules such as keeping quiet, while they will also be challenged by peer students' inappropriate social behaviors. A series of tasks, including looking for a book, sharing seats with others, queuing and checking out a book at the circulation counter, etc., will also be addressed in this scenario. Like scenario 3, scenario 4 recreates another location in the school where social interactions happen frequently - the tuck shop. Children will be challenged in two emotion cognition situations in scenario 4, one being their chosen snack or food has been sold out and the other being a student jumping the queue at the cashier. In scenario 1 to 4, task checklists will appear in the scene to guide the children through the scenarios and visual clues are provided as prompts and aids. Normally, each of the scenario 1 to 4 will be repeated for 6 times in 3 weeks, with gradually increased number of challenges. Moreover, in the consolidation and generalization scenario, the social skills and coping skills taught in scenario 1 to 4 are tested for generalization in a new scene - having physical education class on the playground. The consolidation and generalization scenario will be used for two sessions. In the first session, the children will be asked to interact with the scenario alone with very limited assistance from the trainer. In the second session, the trainer will help the children to reflect their affective and social responses during the first session. Fig. 2 presents screenshots and actual use of scenario. Details regarding the training procedure and design are described in Section 4.3.

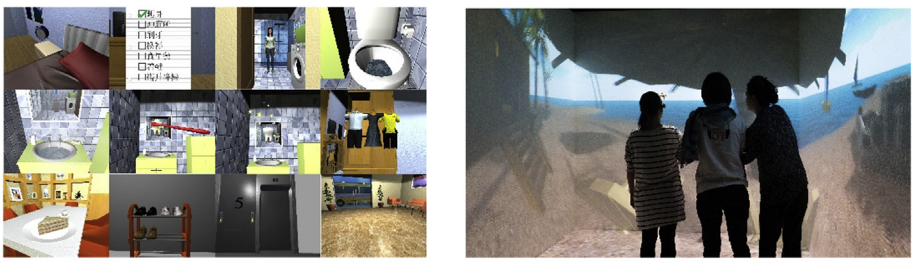
4. Method

4.1. Participants

Primary school-aged children with clinical diagnoses of ASD were eligible for participation in this study. Prior to the actual study, a pilot group was recruited to try out the prototypes in order to fine tune the design. A total of 127 children were referred by 16 different local mainstream primary schools across Hong Kong to join our program. After initial contact with the parents, 3 children were eventually lost to contact, 1 child declined to participate and 29 children were unable to participate at offered time period or their specific time slots. In the end, 94 children (86 boys and 8 girls; mean age = 108.4 months) participated in the study. A complete flow of this process is presented in Fig. 3.

All children and parents were invited to an information session while the parents provided consent for the children to participate in the study. All personal information was not disclosed and they understood that their participation was completely voluntary and that they could withdraw from the study at any time.

These children were divided into two groups, namely Group 1 and Group 2. Group 1 commenced in October 2015, while Group 2 commenced in February 2016. Furthermore, a pilot group of 20 children commenced in July 2015 and served to test out the design and the instructional plan of the scenarios via experiencing the early prototypes of the system. Group 1 and Group 2 participants were invited to complete the training with all the learning scenarios and provided data for the evaluation of the effectiveness of the program. There were 47 participants in each group respectively, with one dropout in Group 1 and two dropouts in Group 2. There were 42 male participants and 5 female participants in Group 1, and 44 male participants and 3 female participants in Group 2. The mean age for Group 1 was 106.3 months while the mean age for Group 2 was 110.5 months. Table 1 gives the demographics of the participants.



(a) (b)

Fig. 2. Virtual reality content. (a) screenshots of scenario 1; (b) emotion control and relaxation scenario in use.

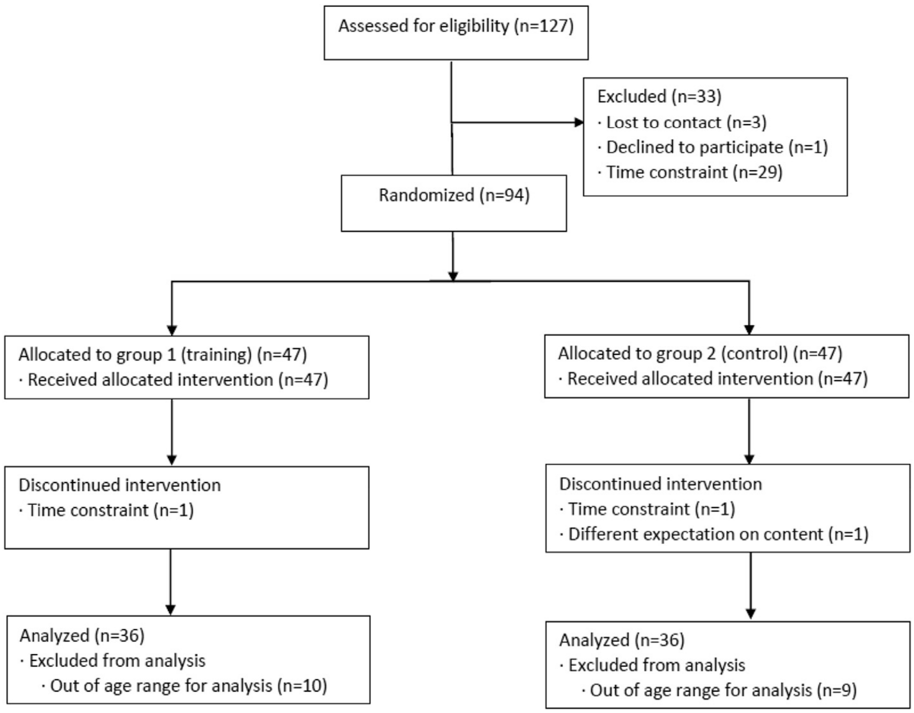


Fig. 3. Flow diagram of the study.

Table 1
Demographics of the vPAD participants.

	No. of Participant	No. of Dropouts	Gender, male (%)	Age[a], mean (SD)
Group 1	47	1	42 (89.4)	106.3 (17.12)
Group 2	47	2	44 (93.6)	110.5 (20.22)
Total	94	3	86 (91.5)	108.4 (18.75)

[a] Age at beginning of training in months.

For analysis purposes, we have used age as a filter and only included those from 7 years old through 10 years old (84–131 months). As such, there were 36 participants in the training group (Group 1) and 36 in the control group (Group 2).

4.2. Measures

Participants were asked to complete pre- and post-assessments before and after training respectively. In order to compare between participants receiving training and those not receiving training, group 1 was treated as a training group while group 2 was treated as a control group. Therefore, all children from group 1 and 2 completed pre-assessments in October 2015 and post-assessments in February 2016. During October 2015 to February 2016, participants of group 1 received the VR-enabled training, while participants of group 2 did not.

Before training, participants completed Raven's Progressive Matrices (RPM) (Raven & Court, 1998) while their parents completed Childhood Autism Spectrum Test (CAST) (Williams et al., 2005). Before and after training, participants completed measures including Faces Test (Baron-Cohen, Wheelwright, & Jolliffe, 1997) with adaptations and Eyes Test (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001) to test for emotion recognition and Psychoeducational Profile, Third Edition (PEP-3) for emotion expression, regulation and social reciprocity. (Schopler, Lansing, Reichler, & Marcus, 2004). Their parents completed Adaptive Behavior Assessment System, Second Edition (ABAS-II) to report on their social adaptive skills (Harrison & Oakland, 2003).

The RPM are tests for measuring nonverbal ability and come in different versions (Raven & Court, 1998). The standard progressive matrices (SPM) consist of 60 patterned diagrams with a part missing to be identified. The SPM are norm-referenced in Hong Kong.

The CAST was previously known as the Childhood Asperger Syndrome Test (Williams et al., 2005). It is a 37-item parental questionnaire to screen for autism spectrum disorders. The cut-off score is 15 for possible ASD or related social-communication difficulties. As our participants were required to provide documentation of ASD diagnosis before entering the study, the CAST was only used for the researchers to better understand the profile of the participants.

The Faces Test is a 20-item test with a western female photos modeling simple emotions and complex mental states (Baron-Cohen et al., 1997). Since our study is conducted in Hong Kong where participants live among Chinese population, a Chinese model was invited to pose the same facial expressions. The photos were taken under standardized lighting with the same emotions as the original version. Only the best pictures that closely replicate the emotions were chosen by two independent researchers. The photographs were displayed digitally with the emotion or mental state word underneath the photo. Although a Chinese translation was available online, they were not commonly used Chinese words. Therefore, additional translation of those emotions was done by independent translators and all discrepancies were resolved by consensus. Participants would learn directly about these emotions and mental states in two of the scenarios. The test was also delivered to the general population to test and ensure that the emotions were easily recognized.

The Eyes Test is a 28-item test with photos of different people's eyes (Baron-Cohen et al., 2001). The test was designed to examine social intelligence by studying how well an individual could identify the photographed person's feelings or thoughts. This task is relatively more difficult than the Faces Test, given that participants have to choose the correct emotions from four choices, the feelings and thoughts are more complex, and that only the eyes region of a person's face would be shown. Participants would not be exposed to these learning contents during the training.

The PEP-3 assesses skills and behaviors in children with ASD (Schopler et al., 2004). The test consists of ten performance subtests but only social reciprocity and affective expression subtests were administered as these two subtests measure our objective skills. The test is originally designed for children with developmental age up to 7 years old. Some of the participants may have a higher developmental age but it was utilized as one of the measures under norm referencing. The Psychoeducational Profile-Revised (PEP-R) was translated into Chinese and a validation study was done in Hong Kong to examine the psychometric properties (Shek, Tsang, Lam, Tang, & Cheung, 2005). The current Chinese PEP-3 edition was also well validated and was used in this study (Shek & Yu, 2014).

The ABAS-II is a norm-referenced assessment of adaptive skills and functioning (Harrison & Oakland, 2003). Parents completed five subscales in the skill areas of communication, community use, leisure, self-direction and social. The official version from Taiwan was used while some minor adaptations on wordings were made in order to match up with the actual social environment in Hong Kong.

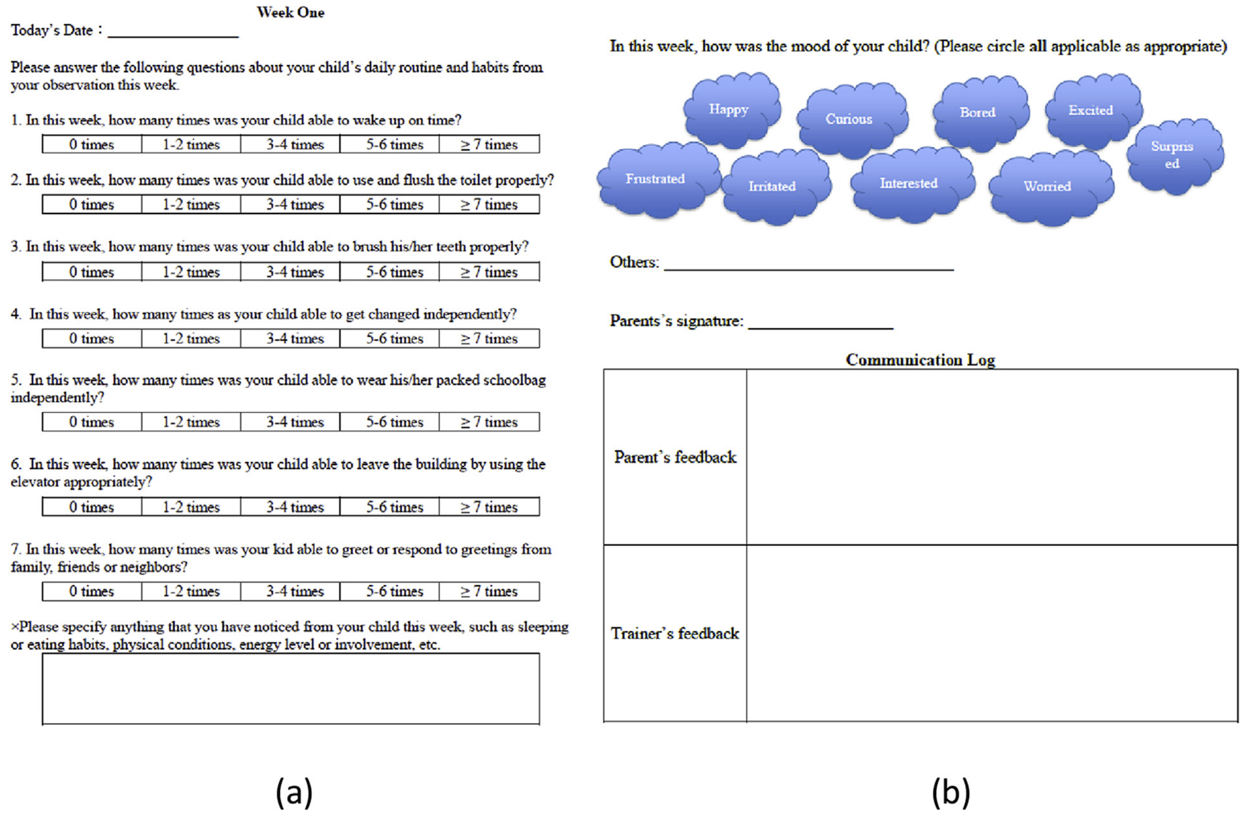
The data collected from the questionnaires and assessments were recorded into SPSS. Subscale scores were calculated for PEP-3 and ABAS-II while total scores were calculated for the other tests. The raw scores from RPM were converted into approximates of non-verbal IQ scores while taking age into consideration.

4.3. Methodology and training procedure

Each of the participants joined the study after the consent and the child's medical report have been received. The child would first join the pre-assessment, which was carried out within a window of two weeks before the first training session started. The training lasted for 14 weeks, and there were two sessions in each week, making a total number of 28 sessions. After the child finished the training, he or she would join the post-assessment, which was carried out within a window of two weeks after the last training session finished.

Each of the 28 training sessions consisted of three stages, i.e. briefing, VR-enabled training, and debriefing. Typically, 3 to 4 children with similar ages joined the sessions together. During briefing, the trainer and the assistant worked with the children to get them both cognitively and affectively ready for the VR-enabled training. The trainer also assessed the conditions of the children during the briefing. The VR-enabled training adopted the group therapy approach. The children in the same training session interacted with the virtual environment in turns. Specifically, the trainer would guide one child in the half-CAVE virtual reality environment, while the others observed at the back. The assistant facilitated those children who were observing and helped them to concentrate and complete a worksheet. The worksheet contained a series of questions regarding the virtual reality learning scenarios. The questions on the worksheet could only be answered if the children observed how their peers interacted with the scenario. This design aimed to encourage the observers to be focused and concentrated during observation and to think and reflect on others' responses to the simulated scenes with social occasions. The virtual reality enabled training would usually last for 40 min. Specifically, each child received an average of 10 min of direct exposure to the virtual reality environment and 30 min of observing. After that the VR-enabled training, the children would be debriefed based on the virtual reality learning scenario they had just experienced and observed in order to support for generalization and to clarify any miscomprehension they may have on the simulated scenes with social occasions.

The uniqueness of this program, when being compared to other VR-enabled approaches, is that it focuses not only on the VR-enabled training, but also supports in other settings beyond virtual reality. To this end, the trainers have developed a set of school teacher and parent training materials to help those who are close to the children provide necessary supports in both school and home setting. Both the school teachers and the parents were invited to observe the training on two high resolution monitors, which display real-time video captured from two different angles. A parent log book was also designed and distributed to the parents so that they could observe, rate, assess their children in the home setting, as well as provided feedbacks to the trainers. An example page of the parent logbook is shown in Fig. 4. Moreover, the school teachers were briefed on the objectives of this program, the contents of the virtual reality scenes and some basic intervention strategies, so that they could provide necessary supports to the children in the school setting.


Week One

Today's Date : ______________

Please answer the following questions about your child's daily routine and habits from your observation this week.

1. In this week, how many times was your child able to wake up on time?

0 times	1-2 times	3-4 times	5-6 times	≥ 7 times

2. In this week, how many times was your child able to use and flush the toilet properly?

0 times	1-2 times	3-4 times	5-6 times	≥ 7 times

3. In this week, how many times was your child able to brush his/her teeth properly?

0 times	1-2 times	3-4 times	5-6 times	≥ 7 times

4. In this week, how many times as your child able to get changed independently?

0 times	1-2 times	3-4 times	5-6 times	≥ 7 times

5. In this week, how many times was your child able to wear his/her packed schoolbag independently?

0 times	1-2 times	3-4 times	5-6 times	≥ 7 times

6. In this week, how many times was your child able to leave the building by using the elevator appropriately?

0 times	1-2 times	3-4 times	5-6 times	≥ 7 times

7. In this week, how many times was your kid able to greet or respond to greetings from family, friends or neighbors?

0 times	1-2 times	3-4 times	5-6 times	≥ 7 times

×Please specify anything that you have noticed from your child this week, such as sleeping or eating habits, physical conditions, energy level or involvement, etc.

In this week, how was the mood of your child? (Please circle all applicable as appropriate)

Happy Curious Bored Excited Surprised Frustrated Irritated Interested Worried

Others: ______________

Parents's signature: ______________

Communication Log

Parent's feedback	
Trainer's feedback	

(a) (b)

Fig. 4. Parent log book. (a) parents' feedbacks on behavioral improvements related to the virtual reality scenario; (b) parents' feedbacks on children's emotion status.

4.4. Data analyses

All data analyses were performed with IBM SPSS Statistics 23. Descriptive statistics were calculated for the demographics and clinical characteristics of each group. To compare the demographics and baseline data between the two groups, independent-sample t-tests were used for means and Chi-square tests were used for frequencies. Mixed repeated measures ANOVA was used to determine if there is an interaction between the difference in scores before and after training across two groups. A significance level of 0.05 was used across all analyses.

5. Results

5.1. Preliminary analyses

After filtering for age, 72 subjects were included in the analyses. There were 64 boys (88.9%) and 8 girls (11.1%). The age of the subjects ranged from 85 to 131 months, with a mean age of 106.3 months (SD 13.53). All subjects completed Raven's Test and the mean score was 95 (SD 17.79), representing a population of normal intelligence. Subjects also completed CAST and the mean score was 19.1 (SD 5.51), which is above the cutoff for possible ASD or related social-communication difficulties. A total of 36 subjects participated in the training group. All of them have completed the program. There were 31 boys (86.1%) and 5 girls (13.9%). The mean age was 107.6 months (SD 13.27). The mean Raven's score was 94.9 (SD 17.49) and the mean CAST score was 19.7 (5.56). Another 36 subjects were in the control group and had completed the same assessments without receiving training in that time period. There were 33 boys (91.7%) and 3 girls (8.3%). The mean age was 104.8 months (SD 13.83). The mean Raven's score was 95.1 (SD 18.34) and the CAST score was 18.4 (5.47).

A summary of descriptive analyses on demographics and baseline measures of the two training and control group are presented in Table 2. There was no significant difference on demographics and baseline measures between the two groups at the beginning of the training. The CAST scores reflect high level of ASD symptoms which reflect their documented ASD diagnoses, while the Raven's score suggest a normal intelligence among our subjects which also corresponds to their placement in mainstream schools.

Table 2
Descriptive statistics of the participants included in analyses.

	Training (n = 36)	Control (n = 36)	Training vs. Control
	n (%)	*n (%)*	*p*
Sex (male)	31 (86.1)	33 (91.7)	0.453
	M (SD)	*M (SD)*	
Age (months)	107.6 (13.27)	104.8 (13.83)	0.769
Raven's score	94.9 (17.49)	95.1 (18.34)	0.465
CAST score	19.7 (5.56)	18.4 (5.47)	0.684

Paired sample t-tests were performed to look at the difference in scores in each group. As our primary outcomes, children from training group scored higher on emotion expression and regulation after the training (M = 20.2, SD = 3.00) than before the training (M = 18.9, SD = 3.57, t(35) = −2.174, p = 0.037). For the other primary outcome on social interaction and adaptation, children from the training group scored higher after the training (M = 21.8, SD = 2.99) compared to before training (M = 20.2, SD = 3.43, t(35) = −3.987, p < 0.000). For secondary outcomes, children from the training group improved on emotion recognition after training although the results were statistically insignificant. The results for adaptive skills also showed improvements. But the improvements were statistically insignificant. For the control group, all measures remained insignificant except interestingly, data has shown improvement on two parent-reported ABAS subscales. Children scored higher on communication in the post-test (M = 47.5, SD = 14.45) than in the pre-test (M = 42.9, SD = 11.93, t(35) = −2.273, p = 0.029). They also scored higher on community use in the post-test (M = 33.4, SD = 20.16) than in the pre-test (M = 26.7, SD = 15.94, t (35) = −2.148, p = 0.039). A summary of mean scores across different measures for both training and control group is summarized in Table 3.

5.2. Primary analyses

Mixed repeated measures ANOVA was used to determine if there is an interaction between the difference in scores before and after training across two groups. It was computed with treatment group as the between subject factor and each of the measures as within-subject variables respectively. A summary of the results is presented in Table 4. There was a statistically significant interaction between the group and time on affective expressions, $F(1, 70) = 5.223$, $p = 0.025$, partial $\eta 2 = 0.069$. There was also a significant interaction between group and time on social reciprocity, $F(1, 70) = 7.769$, $p = 0.007$, partial $\eta 2 = 0.100$. The profile plots of the estimated marginal means on these two analyses are presented in Fig. 5.

The results for secondary outcomes were insignificant. For emotion recognition, there was no statistically significant interaction between group and time on Faces Test, $F(1, 70) = 0.188$, $p = 0.666$, partial $\eta 2 = 0.003$. Similarly, there was no statistically significant interaction between group and time on Eyes test, $F(1, 70) = 0.470$, $p = 0.495$, partial $\eta 2 = 0.007$. For adaptive skills, there was no statistically significant interaction between group and time on communication, $F(1, 68) = 0.291$, $p = 0.592$, partial $\eta 2 = 0.004$. There was statistically significant interaction between group and time on communication, $F(1, 68) = 7.033$, $p = 0.010$, partial $\eta 2 = 0.094$, but control group had improved instead of the training group. There was no statistically significant interaction between group and time on leisure, $F(1, 68) = 0.539$, $p = 0.465$, partial $\eta 2 = 0.008$. There was no statistically significant interaction between group and time on self-direction, $F(1, 68) = 3.897$, $p = 0.052$, partial $\eta 2 = 0.054$. Finally, there was also no statistically significant interaction between group and time on social, $F(1, 68) = 1.000$,

Table 3
Comparison of main measures within groups.

Measures	Training (n = 36)					Control (n = 36)				
	Pre-training	Post-training				Pre-training	Post-training			
	M (SD)	*M (SD)*	*d*	*t*	*p*	*M (SD)*	*M (SD)*	*d*	*t*	*p*
Emotion expression & regulation										
PEP-3 Affective Expressions	18.9 (3.57)	20.2 (3.00)	0.394	−2.174	0.037*	17.0 (4.01)	16.6 (4.85)	0.090	0.909	0.370
Social interaction										
PEP-3 Social Reciprocity	20.2 (3.43)	21.8 (2.99)	0.497	−3.987	0.000***	19.6 (4.14)	19.4 (4.15)	0.048	0.293	0.771
Emotion recognition										
Faces test	16.6 (2.36)	17.2 (2.33)	0.255	−1.233	0.226	15.4 (3.01)	15.7 (2.90)	0.102	−0.721	0.476
Eyes test	12.8 (3.54)	13.5 (3.98)	0.186	−1.374	0.178	12.1 (3.61)	12.3 (3.70)	0.055	−0.328	0.745
Adaptive Skills										
ABAS Communication	44.1 (14.38)	47.2 (11.75)	0.236	−1.866	0.071	42.9 (11.93)	47.5 (14.45)	0.347	−2.273	0.029*
ABAS Community Use	36.7 (16.22)	33.0 (15.50)	0.233	1.594	0.120	26.7 (15.94)	33.4 (20.16)	0.369	−2.148	0.039*
ABAS Leisure	40.9 (14.07)	42.4 (22.55)	0.080	−0.453	0.654	35.5 (10.80)	40.1 (13.41)	0.379	−1.881	0.068
ABAS Self Direction	40.1 (17.01)	38.2 (16.35)	0.114	0.813	0.422	34.5 (15.18)	40.4 (19.80)	0.335	−1.878	0.069
ABAS Social	41.8 (15.19)	41.8 (14.62)	0.000	−0.014	0.989	38.0 (12.00)	41.2 (14.62)	0.239	−1.387	0.174

*p < 0.05; **p < 0.01; ***p < 0.001.

Table 4
Results on mixed repeated measures ANOVA.

Measures	df (between)	df (within)	F	p	partial η2
Emotion expression and regulation					
PEP-3 Affective Expressions	1	70	5.223	0.025*	0.069
Social interaction					
PEP-3 Social Reciprocity	1	70	7.769	0.007**	0.100
Emotion recognition					
Faces Test	1	70	0.188	0.666	0.003
Eyes Test	1	70	0.470	0.495	0.007
Adaptive skills					
ABAS Communication	1	68	0.291	0.592	0.004
ABAS Community Use	1	68	7.033	0.010**	0.094
ABAS Leisure	1	68	0.539	0.465	0.008
ABAS Self Direction	1	68	3.897	0.052	0.054
ABAS Social	1	68	1.000	0.321	0.014

*p < 0.05; **p < 0.01

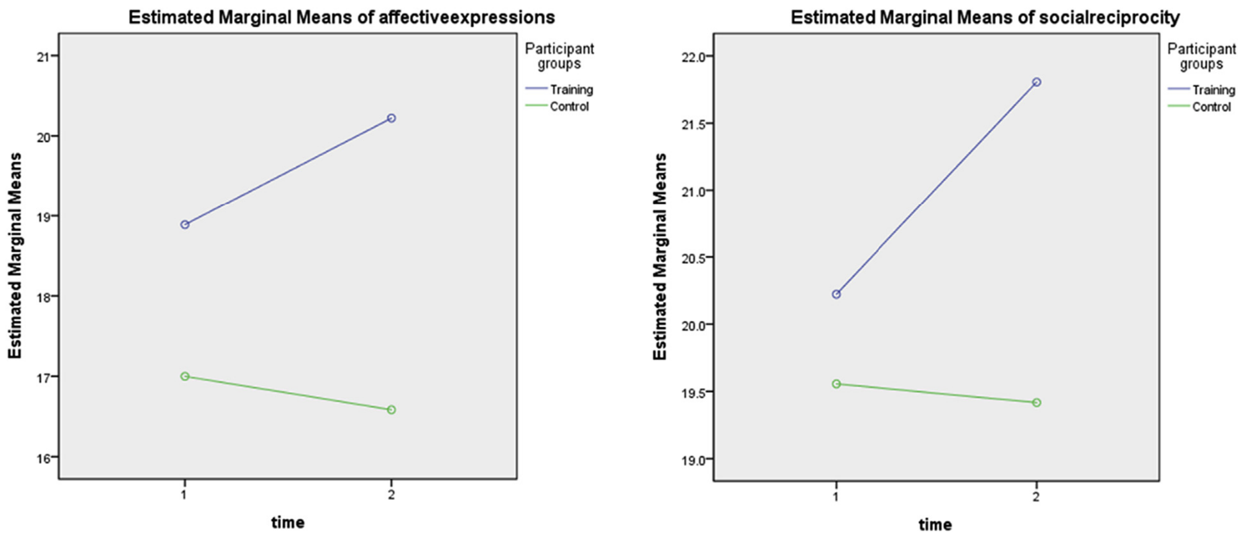


Fig. 5. Profile plots on all primary outcomes: affective expressions (left), social reciprocity (right).

$p = 0.321$, partial $\eta 2 = 0.014$. The profile plots of the estimated marginal means on these analyses are presented in Figs. 6 and 7.

6. Discussion

This study examined the efficacy of a virtual reality-based training program on emotional and social adaptive skills. We hypothesized that children with ASD would improve on their emotional and social adaptation skills after the training, particularly on emotion expression and regulation, and social interaction. Virtual reality is considered a credible tool in education and has shown increased importance and evidence through the last decade. Regardless of the adaptation of virtual reality technology, children participated in the training with peers and a trainer to maximize social interaction and to bridge generalization from virtual environment to real life situations. The results showed that the 28-session training improved emotion expression, emotion regulation and social interaction substantially. Although the majority of training was completed under virtual environment, observation and assessments were completed in normal room settings, suggesting evidence of generalization in such skills. Despite protocol development procedures and the adoption of waitlist control design resulting in a lower number of participants in analyses, nearly a hundred children with ASD participated in the complete training and a lot of positive feedbacks were received.

Throughout the day, children receive an enormous amount of auditory information from instructions given by parents and teachers. This may cause sensory overload in children easily, while the problem is more prominent in children with ASD. Additionally, a lot of the underlying sensory abnormalities in people with ASD may account for some social awkwardness. For instance, hyposensitivity or hypersensitivity to noise may cause embarrassment in social situations. Our program incorporated concepts and theories of education and psychology to the design of an effective training for this population.

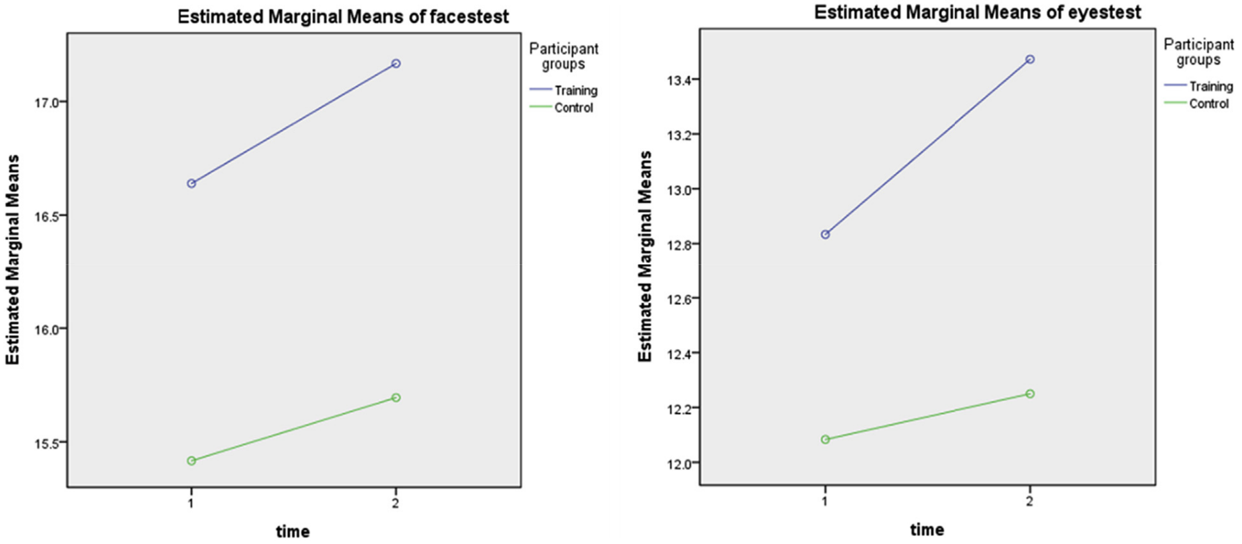


Fig. 6. Profile plots on secondary outcome: Faces Test (left), Eyes Test (right).

Another strength of using virtual reality in social adaptive skills training is that it allows for embarrassing and dangerous social situations without risking the children in real lives (Parsons & Mitchell, 2002). For instance, an important part of community use is safety and road crossing is a big part of it. Our study has enabled the possibility for children to witness inappropriate and dangerous road crossing behaviors and model appropriate responses and behaviors in a safe environment. This is would be unsafe and ethically inappropriate to demonstrate in a real life. Our study also targeted on the repetitive characteristic of children with ASD and allowed for repeated exposure and practice, in which it is not as nearly as easy to set up the same exact scenarios for children repeatedly in the real environment.

In the presented study, there is strong evidence showing significant improvement on our primary outcomes of emotion expression, regulation and social interaction. However, despite observed in-class improvement, the children did not improve significantly on emotion recognition measures including Faces Test and Eyes Test. One explanation to the lack of significant results would be the narrow score range in some of the tasks. As children have different capabilities, some children scored some near perfect scores in pre-assessment and there was little room left for improvement on these assessment tools. Unfortunately, it would be very unlikely to have an assessment tool that is of similar difficulty for all participants in a study despite much effort in matching participants' characteristics. On the other side, a critic of the Faces Test and Eyes Test is that the stimuli are static while interactions in real life are almost never static (Baron-Cohen et al., 2001). Children with ASD have difficulties with subtle cues, hence a visual presentation of emotion may not be representable of the whole picture. This may also explain why improvement observed in social interactions in class is not observed in the standardized measures. Nevertheless, measures observed and rated by our raters, including affective expression and social reciprocity, match up more closely with what was observed in the virtual reality learning environment.

On comparing the adaptive skills across two groups, it was surprising to see that the control group had improved significantly more than the training group. Parents have sent us feedback on how the training has helped their children in their social adaptive skills, yet it was also surprising to see insignificant result on this measure. However, the scale is all based-on parent-report responses and is subjected to response bias. Furthermore, because of the waitlist design, the control group completed the "post-assessment" before their actual training. Naturally, parents would often wish for the best for their children, so they may have the tendency to respond in favor of their children. Although the outcome of assessment has no effect on the intake or quality of training, parents in Hong Kong are often predisposed to such cultural beliefs that all assessments are rooted in a competitive nature. Therefore, there may be social desirability bias as the parents may want their children to look good or score for a better chance before the start of their training. On the other hand, a lot of emotional and social adaptive skills require close observation of behaviors. Given the competitive and busy lifestyle in Hong Kong, it is not uncommon for both parents to be working while a domestic helper or relative is responsible with taking care of the children. Parents may base their responses on their expectation rather than daily observation. This may result in discrepancies in the questionnaires. Similarly, as time pass along with the communication between the trainer and the parents, the parents may gain better understanding on their child after the training. Therefore, they may view the descriptions and the scales of the instruments differently than before.

We were not able to accept several participants referred to us by schools to join the vPAD program. The main reason is the inability to accommodate individual participants' desired time preference due to the study nature and program design. Flexibility may be enhanced if the class frequency is reduced to once a week but subsequent effects remain unknown and

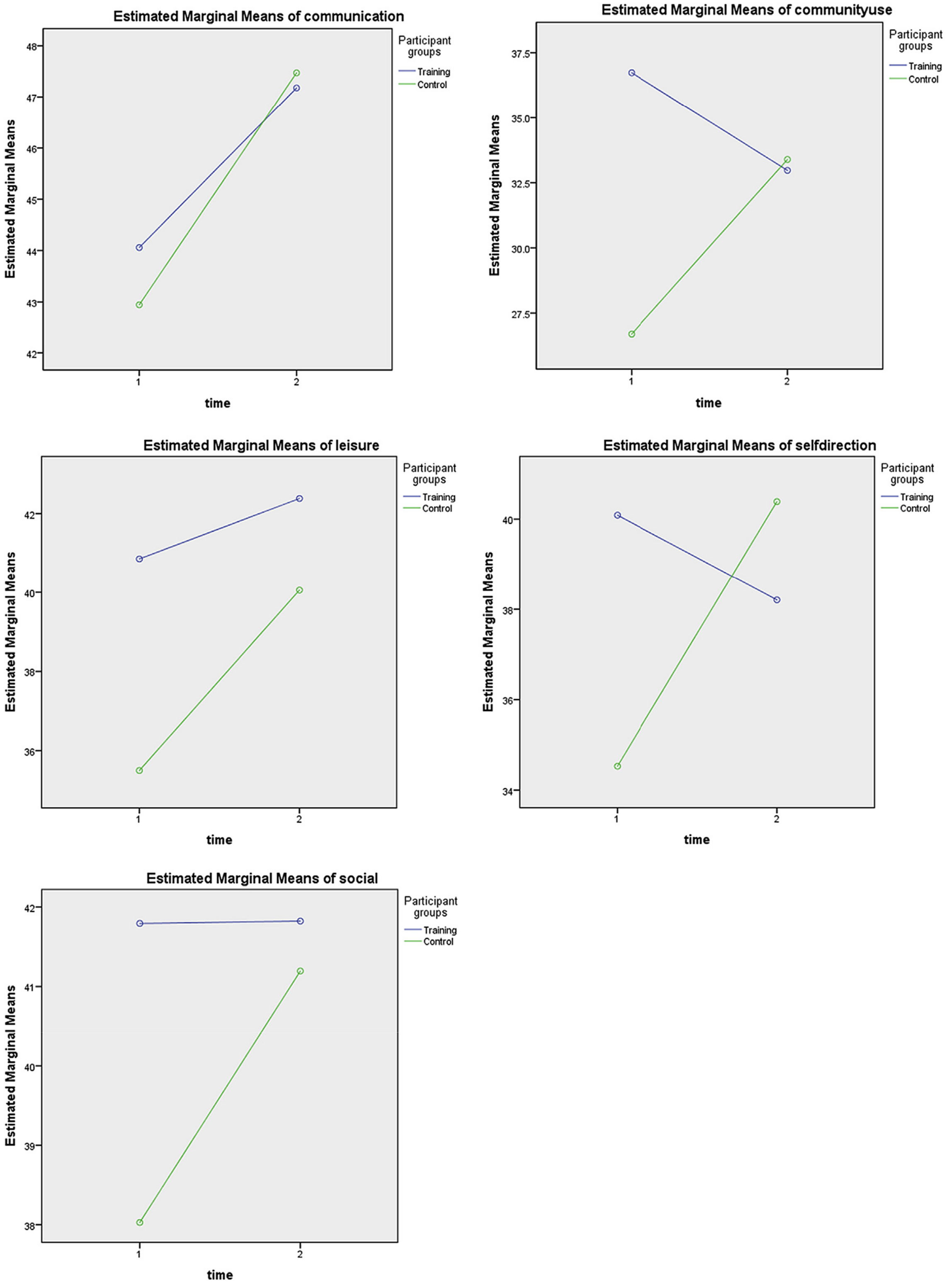


Fig. 7. Profile plots on secondary outcomes on adaptive skills: communication (first left), community use (first right), leisure (second left), self-direction (second right) and social (bottom left).

there may be a higher dropout with a prolonged period of program (approximately 14 weeks versus 28 weeks). Flexibility may also be enhanced if the training is delivered to one or two children at a time, instead or three or four in the current study. However, this setup may limit the amount of social contact and the emphasis of the training may be shifted to emotion expression, while some extra group sessions may be required for practices.

The study has demonstrated feasibility and effectiveness of virtual reality-based emotional and social adaptive skills training in children with ASD and normal intelligence. The current vPAD program hardware is located in our Centre and is expensive to be widely implemented across a larger area to better serve the population. A dilemma on using a relatively mobile device such as a 3D monitor would be the much lessened immersiveness and the subsequent uncertainty on effectiveness. Future research should focus on better polishing the details on such training program as well as using and comparing different modes of virtual reality-based environment. Furthermore, it would be interesting to conduct and compare trainings of different intensity and lengths and on its relative effectiveness. It would be best to conduct long-term and large-scale studies although this is always a difficulty especially when working with this specific population.

7. Conclusion

This study has demonstrated clear feasibility in using virtual reality for the training of children with ASD. The findings from this study provide clear support in improving social skills and emotions expression and regulation for children with ASD. This demonstrates that virtual reality is a promising medium in delivering fun and motivating learning or training program for children with ASD. Future study may consider conducting training programs of a different length or intensity to examine its related effectiveness. Other forms of virtual reality learning environment should also be examined to maximize cost-effectiveness and its target population. Additional research with larger groups is needed to better examine the effectiveness of virtual reality-based learning programs for children with ASD.

Acknowledgement

This research is supported by the HKSAR Quality Education Fund (QEF) (Project No.: 2013/0223).

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Pub. http://doi.org/10.1176/appi.books.9780890425596.744053

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42*(2), 241–251.

Baron-Cohen, S., Wheelwright, S., & Jolliffe, A. T. (1997). Is there a" language of the eyes"? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual Cognition, 4*(3), 311–331.

Bekele, E., Crittendon, J., Zheng, Z., Swanson, A., Weitlauf, A., Warren, Z., et al. (2014). Assessing the utility of a virtual environment for enhancing facial affect recognition in adolescents with autism. *Journal of Autism and Developmental Disorders, 44*(7), 1641–1650. http://doi.org/10.1007/s10803-014-2035-8.

Cai, Y., Chia, N. K. H., Thalmann, D., Kee, N. K. N., Zheng, J., & Thalmann, N. M. (2013). Design and development of a virtual dolphinarium for children with autism. *IEEE Transactions on Neural Systems and Rehabilitation Engineering : A Publication of the IEEE Engineering in Medicine and Biology Society, 21*(2), 208–217. http://doi.org/10.1109/TNSRE.2013.2240700.

Cruz-Neira, C., Sandin, D. J., & DeFanti, T. A. (1993). Surround-screen projection-based virtual reality. In *Proceedings of the 20th annual conference on computer graphics and interactive techniques - SIGGRAPH '93* (pp. 135–142). http://doi.org/10.1145/166117.166134.

Cruz-Neira, C., Sandin, D. J., DeFanti, T. a., Kenyon, R. V., & Hart, J. C. (1992). The CAVE: Audio visual experience automatic virtual environment. *Communications of the ACM, 35*(6), 64–72. http://doi.org/10.1145/129888.129892.

Dawson, G., & Watling, R. (2000). Interventions to facilitate auditory, visual, and motor integration: A review of the evidence. *Journal of Autism and Developmental Disorders, 30*, 415–421.

DeFanti, T. A., Dawe, G., Sandin, D. J., Schulze, J. P., Otto, P., Girado, J., et al. (2009). The StarCAVE, a third-generation CAVE and virtual reality OptIPortal. *Future Generation Computer Systems, 25*(2), 169–178. http://doi.org/10.1016/j.future.2008.07.015.

Foxx, R. M. (2008). Applied behavior analysis treatment of autism: The state of the art. *Child and Adolescent Psychiatric Clinics of North America, 17*(4), 821–834.

Gray, C. A., & Garand, J. D. (1993). Social stories: Improving responses of students with autism with accurate social information. *Focus on Autistic Behavior, 8*(1), 1–10.

Harrison, P. L., & Oakland, T. (2003). *Adaptive behavior assessment system* (2nd ed.). San Antonio, TX: The Psychological Corporation.

Higgins, K., & Boone, R. (1996). Creating individualized computer-assisted instruction for students with autism using multi-media authoring software. *Focus on Autism and Other Developmental Disabilities, 11*(2), 69–78.

Hilton, C., Graver, K., & LaVesser, P. (2007). Relationship between social competence and sensory processing in children with high functioning autism spectrum disorders. *Research in Autism Spectrum Disorders, 1*(2), 164–173.

Ip, H. H. S., Hay, Y., & Tang, A. C. C. (2002). Body-brush. In *Proceedings of the tenth ACM international conference on multimedia - MULTIMEDIA '02* (p. 664). http://doi.org/10.1145/641007.641144.

Jacobson, N. S. (1978). Specific and nonspecific factors in the effectiveness of a behavioral approach to the treatment of marital discord. *Journal of Consulting and Clinical Psychology, 46*(3), 442–452. http://doi.org/10.1037//0022-006X.46.3.442.

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*(1), 34–44. http://doi.org/10.1007/s10803-012-1544-6.

Lorenzo, G., Pomares, J., & Lledó, A. (2013). Inclusion of immersive virtual learning environments and visual control systems to support the learning of students with Asperger syndrome. *Computers & Education, 62*, 88–101. http://doi.org/10.1016/j.compedu.2012.10.028.

Matsentidou, S., & Poullis, C. (2014). Immersive visualizations in a VR cave environment for the training and enhancement of social skills for children with autism. In *VISAPP 2014-proceedings of the 9th international conference on computer vision theory and applications* (Vol. 3, pp. 230–236).

Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation, 41*, 85–139.

McCann, R. a., Armstrong, C. M., Skopp, N. a., Edwards-Stewart, A., Smolenski, D. J., June, J. D., et al. (2014). Virtual reality exposure therapy for the treatment of anxiety disorders: An evaluation of research quality. *Journal of Anxiety Disorders, 28*(6), 625–631. http://doi.org/10.1016/j.janxdis.2014.05.010.

Merhi, O., Faugloire, E., Flanagan, M., & Stoffregen, T. A. (2007). Motion sickness, console video games, and head-mounted displays. *Human Factors The Journal of the Human Factors and Ergonomics Society, 49*(5), 920–934. http://doi.org/10.1518/001872007X230262.Copyright.

Mesibov, G. B., Shea, V., & Schopler, E. (2005). *The TEACCH Approach to autism spectrum disorders*. Springer Science & Business Media.

O'neill, M., & Jones, R. S. (1997). Sensory-perceptual abnormalities in autism: A case for more research? *Journal of Autism and Developmental Disorders, 27*(3), 283–293.

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*(March 2015), 355–366. http://doi.org/10.1080/08856257.2011.593831.

Parsons, S., Leonard, A., & Mitchell, P. (2006). Virtual environments for social skills training: Comments from two adolescents with autistic spectrum disorder. *Computers and Education, 47*(2), 186–206. http://doi.org/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*(5), 430–443. http://doi.org/10.1046/j.1365-2788.2002.00425.x.

Raven, J. C., & Court, J. H. (1998). *Raven's progressive matrices and vocabulary scales*. Oxford, UK: Oxford Psychologists Press.

Ron, C. W. K., Cheng, S. H., Ip, H. H. S., & Joseph, S. L. K. (2011). Design of affectively evocative smart ambient media for learning. *Computers & Education, 56*(1), 101–111.

Rothbaum, B. O., Anderson, P., Zimand, E., Hodges, L., Lang, D., & Wilson, J. (2006). Virtual reality exposure therapy and standard (in vivo) exposure therapy in the treatment of fear of flying. *Behavior Therapy, 37*(1), 80–90. http://doi.org/10.1016/j.beth.2005.04.004.

Schopler, E., Lansing, M. D., Reichler, R. J., & Marcus, L. M. (2004). *Psychoeducational profile Third edition (PEP-3). Pro-ed, USA*.

Sharples, S., Cobb, S., Moody, A., & Wilson, J. R. (2008). Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays, 29*(2), 58–69. http://doi.org/10.1016/j.displa.2007.09.005.

Shek, D. T. L., Tsang, S. K. M., Lam, L. L., Tang, F. L. Y., & Cheung, P. M. P. (2005). Psychometric properties of the Chinese version of the psycho-educational profile-revised (CPEP-R). *Journal of Autism and Developmental Disorders, 35*(1), 37–44. https://doi.org/10.1007/s10803-004-1029-3.

Shek, D. T. L., & Yu, L. (2014). Construct validity of the Chinese version of the psycho-educational profile-3rd edition (CPEP-3). *Journal of Autism and Developmental Disorders, 44*(11), 2832–2843.

Shores, R. L. (1987). Overview of research on social interaction: A historical and personal perspective. *Behavioral Disorders, 12*, 233–241.

Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., et al. (2013). Real-time human pose recognition in parts from single depth images. *Commun. ACM, 56*(1), 116–124. http://doi.org/10.1145/2398356.2398381.

Sigman, M., & Ruskin, E. (1999). Continuity and change in the social competence of children with autism, Downs syndrome, and developmental delays. *Monographs of the Society for Research in Child Development, 64*(1), 1–114.

Smith, M. J., Ginger, E. J., Wright, K., Wright, M. A., Taylor, J. L., Humm, L. B., et al. (2014). Virtual reality job interview training in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 44*(10), 2450–2463. http://doi.org/10.1007/s10803-014-2113-y.

Steuer, J. (1992). Defining virtual Reality: Dimensions determining telepresence. *Journal of Communication, 42*(4), 73–93. http://doi.org/10.1111/j.1460-2466.1992.tb00812.x.

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*(6), 651–659. http://doi.org/10.1007/BF02172354.

Taylor, R. M., II, Hudson, T. C., Seeger, A., Weber, H., Juliano, J., & Helser, A. T. (2001, November). VRPN: A device-independent, network-transparent VR peripheral system. In *In Proceedings of the ACM symposium on Virtual reality software and technology* (pp. 55–61). ACM.

Watling, R. L., Deitz, J., & White, O. (2001). Comparison of sensory profile scores of young children with and without autism spectrum disorders. *American Journal of Occupational Therapy, 55*(4), 416–423.

Williams, J., Scott, F., Stott, C., Allison, C., Bolton, P., Baron-Cohen, S., et al. (2005). The CAST (childhood Asperger syndrome test) test accuracy. *Autism, 9*(1), 45–68.