



Inclusion of third-person perspective in CAVE-like immersive 3D virtual reality role-playing games for social reciprocity training of children with an autism spectrum disorder

Wei-Te Tsai¹ · I-Jui Lee² · Chien-Hsu Chen¹

Published online: 2 June 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

The present study aimed to improve the ability of children with autism to recognize emotions correctly. We used our third-person perspective role-playing game (TPP-RPG) method to teach social skills and help develop an improved understanding of the six basic emotions. The experiment was divided into two phases: The first involved working with traditional figure card emotional recognition and the second involved a subject entering a 3D cave automatic virtual environment (CAVE) to engage with interactive games. While the traditional graphic card is a static picture that represents one of the six basic human emotions, the virtual reality of CAVE-like immersive 3D role-playing games enables the use of actual picture scene syntheses plus the animation of 3D characters to express emotions. The participating children were instructed to role-play with (1) three-dimensional (3D) virtual role animations and observe (2) two different real-time switchable role-play animations of themselves and their counterpart socially interacting. This single-subject study was based on multiple-baseline, across-subject design and involved 5 weeks of TPP-RPG training intervention. From this research activity, we found that the role-play performance of all three participants rose substantially during the intervention phase and remained significantly higher in the maintenance phase compared to their baseline levels.

Keywords CAVE-like immersive environments · Autism · Third-person perspective · Virtual Reality CAVE · Role-playing game · Human–computer interaction · Empathy

1 Introduction

1.1 The innate defects in autism

Children with autism are generally considered to have birth defects in social behavior, communication barriers, and

repetitive behavior [3, 8]. Children with autism spectrum disorders (ASDs) often find that social interaction with others is difficult [26]. In particular, they are often not sure how to respond to the feelings, facial expressions, and body movements of other people, and in particular what they should say [14]. The nature of these challenges has been well documented in the literature [3, 14]. Many researchers have determined that reciprocity plays a crucial role in maintaining cooperation in society and that this requires the ability to empathize with others [7, 9], to be aware of emotional and interpersonal cues, and to respond appropriately. It is also clear that effective interaction requires nonverbal communication and nonverbal social skills, without which it is difficult to know how other people feel.

1.2 Autism's defects in social reciprocity behavior

Beyond their social interaction difficulties, children with ASD also experience substantial communication barriers, relatively rigid behavior patterns, and a general inability

✉ I-Jui Lee
ericlee@mail.ntut.edu.tw
Wei-Te Tsai
p38021067@mail.ncku.edu.tw
Chien-Hsu Chen
chenhsu@mail.ncku.edu.tw

¹ Ergonomics and Interaction Design Lab, Department of Industrial Design, National Cheng Kung University, No. 1 University Rd., East District, Tainan, Taiwan

² Ergonomics Design and Technology Lab, Department of Industrial Design, National Taipei University of Technology, 1, Sec. 3, Zhongxiao E. Rd., Taipei 10608, Taiwan

to empathize with others [7, 9]. They often also are clearly insensitive to emotional cues, and have difficulty understanding when and how to appropriately respond in everyday social situations [2, 3]. A particular challenge is that children with ASD have problems distinguishing happy social encounters from angry ones because they have difficulty in understanding whether their counterparts are smiling, waving hello, or shaking an angry fist at them [2]. Fortunately, these social cues can be learned through training [24, 27]. For example, role-play and videos are often used to strengthen the interaction skills of children with ASD when learning and socializing [6, 42]. These two methods are commonly used in the field of education and are often applied in the teaching and training of autistic children [24]. However, we consider role-play and watching videos to have a substantial drawback because these methods do not enable the participants to observe themselves in various social situations [2, 18]. Being able to see one is a critical influential factor in training. The ability to observe from the third-person perspective is useful for children because they can use this information to understand and mimic what they see, thereby enhancing their social emotional awareness or self-reciprocity skills [13, 18]. To overcome the limitations of role-play and videos watching, we developed a TPP-RPG training method that enables children to specifically observe and evaluate their own facial expressions, emotions, and body languages from a third-person perspective.

2 Related work

2.1 The training strategies of video modeling commonly used in children with ASD

Current research often uses video modeling (VM) to teach children with ASD to role-play and pretend play [33] so that they can learn to understand the meanings of the facial expressions and body language of their counterparts in various social situations [2, 12, 16, 17, 44]. By watching the videos [6] or undertaking immersion into virtual reality environments [35], children with ASD can often better understand the roles and contexts of the people in the scenarios as well as variations in their body language [12]. Trainers in such settings often ask the children to focus on how the emotions, facial expressions, and body language of the characters in the videos change and then ask them to imitate those movements and expressions [13, 44]. Evidence shows that VM can significantly help children with

ASD to understand and reflect on social behaviors [5, 12, 16, 17, 34, 44].

2.2 The primary training limitation of the VM strategy

However, because VM can teach users only to imitate behaviors and usually does not explain what the behaviors mean or why they should imitate them [5], users often cannot clearly understand why they should learn these behaviors [17]. In addition, a VM experience is almost always presented from the viewer or video creator's perspective, showing a social interaction without establishing a direct connection with the characters in the role-play [31]. In this context, it is unfortunate that VM training does not provide a direct interactive mechanism that allows clear social communication with others, and this frequently makes VM by itself uninteresting to children with ASD [19, 31]. However, children with ASD need to be trained to understand the perspective of those they are interacting with and to see how their own behaviors affect the other person's social reciprocity interactions.

2.3 Applying TPP and RPG methods in CAVE-like immersive 3D virtual reality role-playing games in order to provide social skills training

Because children with ASD lack imagination and do not have the ability to pretend or to play games [33], it is often difficult for them to understand the feelings of others [7, 9] and the meaning of social reciprocity [32]. The past literature has also pointed out that autistic children differ from others in their lack of understanding of visual space in social situations [2, 36]. Having greater skill in this area is especially critical because autism's innate defect is that it diminishes one's capability to understand the feelings of other people and thus the ability to empathize [7, 9] and interact productively with them [43]. Therefore, our research uses a combination of TPP and RPG methods, which overlap images of ASD children in real scenes while allowing the user to interact with others in virtual 3D. Training mechanisms integrated with RPG for self-modeling [12, 30] enable users to view a real-time 3D animation of themselves. This kind of practice makes it possible for children with ASD to simultaneously observe their own facial expressions and body language and those of their counterparts in the RPG situation [17]. We think that the unique perspective obtainable from using a TPP-RPG approach can help children with ASD to see themselves interacting with virtual avatars from the perspective of others and that this makes it possible for them to increase their situational awareness and understanding of overall contexts and social reciprocity.

2.4 Benefits and opportunities of the TPP-RPG training method in social reciprocity training

Beyond situations where CAVE-like immersive 3D virtual reality environments are employed, various techniques and methods have been used to quantify social skills and cognition, with different levels of success [21, 45]. Some group intervention studies have used observational techniques, such as observing conversation styles in simulated scenarios of parties or interviews [27, 41], or the frequency and type of observational interactions have been tracked. The practicality of using virtual reality is due to several unique characteristics which make it particularly useful when attempting to quantify social skills and cognition [4, 35]. Virtual reality represents real life in a safe, controllable way, and repeated exercises and exposures can be used as key elements in treatment [28, 46]. Virtual reality can also provide a natural environment with an infinite number of social scenes and has been proved to replicate social conditions well [45].

In addition, the TPP-RPG training method is effective because it overcomes the reality for autistic children that it is difficult for them to imagine and understand social interactions that they have not previously seen or directly experienced. In contrast to our approach, a traditional training strategy, such as VM, only allows the patient to experience a single virtual character playing with them and does not show them how they themselves are interacting with others. However, if it is possible for them to interact and watch themselves from the perspective of a third person, they become better able to think about their relationship with others because this angle of view allows them to see themselves playing in the virtual context. Our new third-person approach also provides a unique visual framework that helps patients to improve their understanding of abstract social reciprocal relationships in general.

2.5 Gaps beyond the previous studies

Our research, and the results we have been able to achieve, differs from prior work in this area of study in many ways [10, 39]. First, we use a CAVE-like, immersive 3D virtual reality environment to present virtual 3D characters and real participants in ways that overlap. A CAVE-like 3D virtual reality environment can give the participants with ASD an opportunity to focus on their own and others' social interaction. They can observe the response of 3D avatars to their own social interaction gestures and comprehend better whether their self-body social reciprocity is appropriate or not. Being able to see how others respond to them stimulates interest and increases the pleasure of learning for children with ASD. This effect cannot be achieved using conventional videos or VM.

Second, our RPG method allows users to switch their perspective between first- and third-person perspectives. This allows them to overlap observations of the real participant and of the virtual participant. This differs from previous studies which have used augmented reality (AR) technology as a training strategy to help ASD children learn how to model their own facial expression by looking at the 3D augmented facial expressions of another participant's face [18, 31]. AR is limited because it is only able to focus on specific parts of the body [32]. Our method, in contrast, uses a much more engaging whole-body interactive framework. Earlier research has focused on the connections between emotions, facial expressions, and facial characteristics [13, 18, 32], but our research adds to this mix the ability to capture full body language as well. In order to do this, we used two Kinect cameras placed on the front and back of a child and projected that child's image together with the images of 3D characters.

Third, unlike VR, where participants have to wear special glasses [15], our CAVE-like, immersive 3D virtual reality environment allows participants to literally walk into the training environment. This helps children with ASD concentrate and feels a sense of stability. Most VR devices require participants to wear a heavy head-mounted display (HMD) [1]. It has been widely observed that using traditional VR equipment often makes participants feel uncomfortable [40]. The heavy VR equipment can make participants feel as though they are being forced to interact with some kind of ominous medical equipment and often find that it can be uncomfortable to wear. Our system completely eliminates this set of issues.

3 Methods

We developed a multi-perspective (involving both first and third persons), interactive, CAVE-like immersive 3D virtual reality role-playing game for socialization training in greeting others. The aim was to enable children with ASD to observe whether their social reciprocity behaviors were appropriate and correct (Fig. 1).

3.1 Participants

Three children with ASD (three boys, using the pseudonyms Doran, Bob, and Peter) took part in the study. The mean age of this group was 7.7 years, and the age range was 7–9 years. Their intelligence quotient (IQ) scores were: [a] full scale IQ (FIQ): 87; [b] verbal IQ (VIQ): 88; and [c] performance IQ (PIQ): 86 (see Table 1). They were in the same school with numerous typically developing (i.e., non-ASD) students. The inclusion criteria were: (a) a clinical diagnosis of ASD based on DSM-5 criteria [3], (b) affirmed by a physician to have

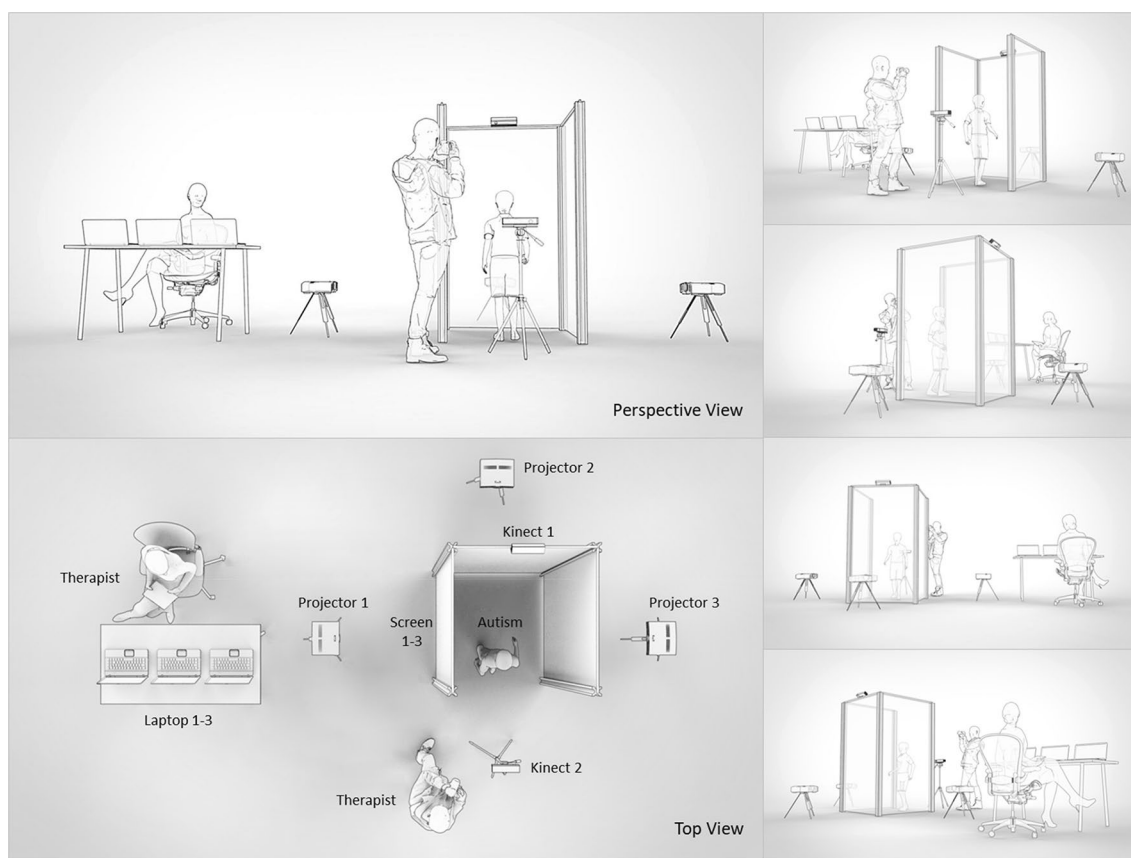


Fig. 1 Users stand in front of the immersive cave screen and interact with the avatars. Two Kinect camera modules are set up: one behind and one in front of the user. In the scene, there is a three-dimensional projection screen (including three projectors), a front-end lens (using

one Kinect camera), a back-end lens (using another Kinect camera), and three laptops, one of which is the main program operation machine. There is also a third camera to record the user's activities with avatars

Table 1 Summarized demographic information of the participants

Participants	Information				
	Age	FIQ	VIQ	PIQ	Diagnosis
Doran	7	72	73	69	ASD
Bob	8	95	96	97	ASD
Peter	8.2	94	95	92	ASD
Mean	7.7	87	88	86	

no comorbidities, (c) not taking prescribed medications, and (d) not undergoing any other therapies at the same time as the testing.

The regular classroom teacher for these three students was responsible for general lessons for all students, but was assisted by a teacher's aide. Once a week, a substitute teacher (an occupational therapist) taught the class. The Vineland II Adaptive Behavior Scales [11] showed that Doran, Bob, and Peter were at a moderate medium adaptive level, which means that they could perform tasks when supervised. During interviews with the parents and therapists of Doran, Bob,

and Peter, we learned that the children thought it was difficult for them to interact and socialize with others. There was a general agreement that they did not understand facial expressions or body language movements in social greetings and therefore often misunderstood the feelings of other people, leading them to shy away from interacting with others.

3.2 Developing the third-person perspective role-playing game training system

In the experiment we undertook, our TPP-RPG training system used two cameras for switching between first-person and third-person perspectives. Using overlapping CAVE-like immersive 3D images of virtual roles, the images produced can be switched among three perspectives: (a) viewing the actions and expressions of the participant from the perspective of a virtual role-player (TPP: shot by front camera); (b) seeing the virtual 3D character from the perspective of the participant (first-person perspective: shot by back camera); and (c) the system can overlap the user's perspective by adding 3D images of a virtual role-player into an interactive

scenario. In addition, a third camera was used to record a panoramic view of the entire training session. This provided footage for subsequent video analysis and discussions between the children and the therapist (Fig. 1).

To conduct the experiment, two Kinect camera modules were set up: one behind and one in front of the user. In the scene, a three-dimensional projection screen (including three projectors) was set up, along with a front-end lens (using one Kinect camera), a back-end lens (using another Kinect camera), and three laptops, one of which was the main program operation machine. A third camera was also included to record the user's interactions with avatars. The construction of the TPP-RPG training system was mainly divided into two categories: hardware and software. The hardware included a projection screen, projectors, two infrared detectors, and a recorder. The software included 3D animation, image synthesis, and an interactive system (Fig. 1).

3.3 Hardware

The projection system was mainly constructed from three white curtains and an aluminum extrusion model (Fig. 1). The primary purpose of the design employed here was to build an immersive environment with a square shape, having a translucent projection screen on three sides, but none on the fourth side. The side where the projection was not installed provided space for the participant to stand. Three projectors were located behind the three screens: One Optoma short-throw projector was used on each side (model K300ST) and behind the back. Where the main screen was located, a NEC high-lumen projector was used. An infrared sensor was used primarily to detect the behavior of the participants and to put the images of the participants into the environment. The instruments used were Microsoft Kinect® devices, designed for developers. The recorder was a digital camera.

3.4 Software

Mainly designed for digital content production, the software that was used for 3D animation in this study was iClone, the image synthesis software was Adobe Photoshop, and the interactive imaging system for AR and VR used the Unity program. A total of six emotional performances were produced in 3D animation, with different versions for boys and girls. In addition, three scenes were designed that were set in parks, living rooms, and classrooms. The scene production background used a 2D photograph, and the foreground was rendered using a portion of the 3D object. Finally, we integrated the Unity system with the Kinect v2 examples and used the MS-SDK and the NuiTrack SDK as a controller with the Kinect.

Consequently, during the experiment, 6 emotions, 12 versions, and 3 scenes were arranged, and the C# program was used to develop random number arrangements and related animation controls.

3.5 Operational scenario and facilities

In our experimental protocol, the following items were embedded inside the system: (a) 6 virtual 3D characters to stimulate interaction between system users and the virtual roles expressing different emotions, showing different facial expressions and using different body language; (b) the three previously described cameras; and (c) a three-sided aluminum support seat and a three-sided projection curtain. (The front side is the main character screen, and the left and right side screens are there to increase immersion.) TPP-RPG system users stand in front of the projection curtain to watch the 3D animations of virtual roles. By overlapping the front and side panels of the screen, the interactive relationship between the user and the virtual role is shown and allows the user to see the three different perspectives provided by the TPP-RPG system (Fig. 2).

3.6 Measurement materials

3.6.1 Social story tests (SSTs)

The scenario scripts (see “Appendix 1”) and related tests designed from multiple-baseline experiments were made by therapists, special education teachers, and parents of children with autism, incorporating the Social Stories™ concept proposed by the researcher Gray [25] for children with ASD. This approach is designed to teach children with ASD to learn social contexts and skills [23, 29, 37]. These SSTs were presented in the form of contextual questions and answers and social story scripts for each augmented greeting behavior fragment were created, and each scenario was associated with a different event. There are 20 questions in the baseline phase, 20 in the intervention phase, and 20 in the maintenance phase, respectively, all with approximately equal difficulty (similar situation and same length). The content validity was confirmed by a panel of experts to determine how well the test items reflected the range of content being measured. After the test, the question answers and expression scores were confirmed by the special educational teacher and the judgment of an expert, and we recorded the answers to determine the correct response rate. In addition, we submitted questionnaires and interviews that related to the results of tests for expert assessment and parental review so as to ensure the social reliability of the tests and to validate that the tests represented situations that were close to real life. In addition, we compared performance before and



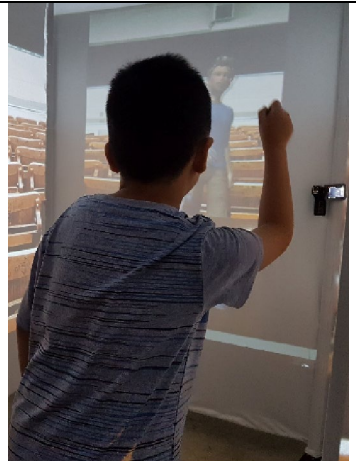
Actual Scenario Situation	Viewing Perspective	Picture Features and Viewing Situations
	Third-person	(1) Seeing one's own body language and facial expressions from the perspective of a virtual character (self-modeling from the interaction)
	Third-person	(2) From the user's perspective (behind) to determine how to socially interact in different situations
	First-person	(3) Observing the physical movement and social situation of the virtual character from the screen

Fig. 2 Three different perspectives

after intervention using the paired-sample t test to determine whether a particular intervention method improved social greeting ability. SPSS 21.0 (SPSS Inc., Chicago, IL) was used for all statistical analyses.

3.7 Role-playing game evaluation

All participants role-played the greeting behavior after SST questions in each session. The therapist evaluated

their greeting behavior on a 5-point Likert scale: 5—excellent, 4—good, 3—fair, 2—poor, and 1—very poor.

Role-playing games (RPG) are a technical training strategy designed to help kids on the autism spectrum build social skills and confidence. They are a great way for parents or teachers to engage with their children at home or school, and for professionals and experienced gamers to use as a group social skills tool. In research, role-playing games have shown that the social interaction required in these games may also develop social or emotional regulation skills in ordinary life [42].

Our study of the greeting behavior of our subjects during role-play took place in three phases (Fig. 3): a baseline phase, an intervention phase, and a maintenance phase. During these phases, we focused on three aspects of participants' greeting behavior: their motoric–physical skills, their social behaviors, and their social cognitive skills. The motoric–physical skills included actions such as smiles, greetings, handshakes, and hugs. The social behaviors included self-introductions, questions and conversations, polite greetings, requests for assistance, expressions of gratitude or praise to others, etc. The social cognitive skills included understanding social clues, obeying rules, empathy, perceptions of feeling, and problem-solving.

4 Experimental conditions

The three children with ASD in our study group have congenital conditions that manifest differently for each of them. Therefore, we used a single-subject research method [37] to confirm the effectiveness of the intervention on individual subjects, despite the fact that the three children were all ASD. To achieve our aim of training children with ASD to recognize facial expressions and body language during social interactions, and to help them respond with appropriate reciprocal behaviors, we created 3D virtual characters that corresponded to their family members, classmates, teachers, friends, etc. We also developed 20 basic and typical social scenarios that our users often find themselves in.

For this study, the parents of each participant signed an informed consent form that was provided by the Ethics Committee of Behavioral and Social Sciences Research, National Taiwan University. This form is used by relevant academic research institutions across the country. Each participating student also signed a youth consent form. Each researcher who undertook experiments utilizing our system had previously been trained in research ethics for more than 6 h and had obtained a certificate to this effect.

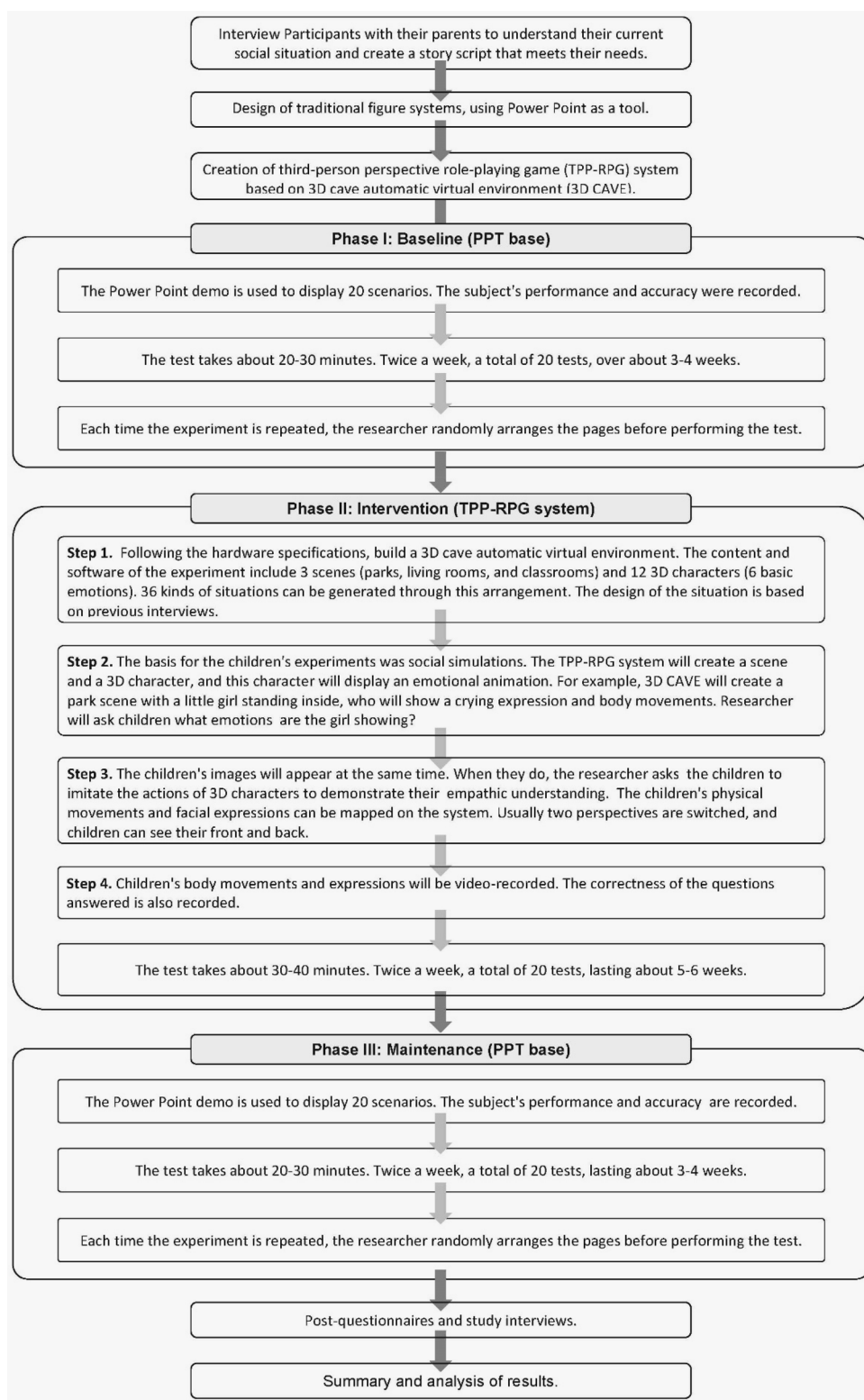
4.1 Baseline phase

In the baseline phase, (a) the therapist talked about and illustrated the facial expressions of the six primary emotional expressions (anger, disgust, fear, happiness, sadness, and surprise) and their corresponding body language; (b) during the instruction, the therapist used slideshows that included the scripts and pictures of each scenario; (c) after the scenario slideshows, the therapist showed pictures of facial expressions with the six basic emotions [22] and body language for each. We followed the Facial Action Coding System (FACS) to define each emotional expression [20] and used input from experts to discuss the body language that corresponds to the six primary emotional expressions [22]; (d) after this section of the training was finished, the therapist asked the children to mimic each facial expression and corresponding body language, recorded it on videotape, and used it as evidence of the TPP-RPG system's effectiveness in subsequent appraisals of results.

4.2 Intervention phase

During the intervention phase (Fig. 4), the ASD children, using our TPP-RPG system, observed and learned about their own and others' facial expressions and body language when interacting socially. We proceeded in seven steps: (a) before the greetings training began, the therapist taught the users how to interact with the TPP-RPG system itself. When the system started, the children's images were mapped on the screen; (b) for each scenario; we provided a script and questions regarding facial expressions and body language. Three scenes were presented: a living room, a classroom, and a park along with 6 basic emotions and 12 3D characters. Two different 3D character animations were presented for each mood; (c) after the animations were presented, the subjects were asked to identify the emotional expressions which they had seen. They were given 6 possible multiple-choice answers representing the 6 basic emotions expressions to choose from; (d) correct answers generated a green circle and switched the user to our TPP mode; (e) the therapist then guided the users as they mimicked the facial expressions and body language they had seen; (f) incorrect answers generated an error signal and images of faces with puzzled expressions; (g) when the test was completed, users were asked to repeat the expressions and body language that they had mimicked and to watch themselves while doing these tasks, and all users individually and separately role-played with the therapist. The settings on the system can switch between two perspectives, which can project the front and back of a participant's body movements. The front allows children to see their facial expressions and compare their expressions with the expressions of the 3D characters on

Fig. 3 Training process for the TPP-RPG system

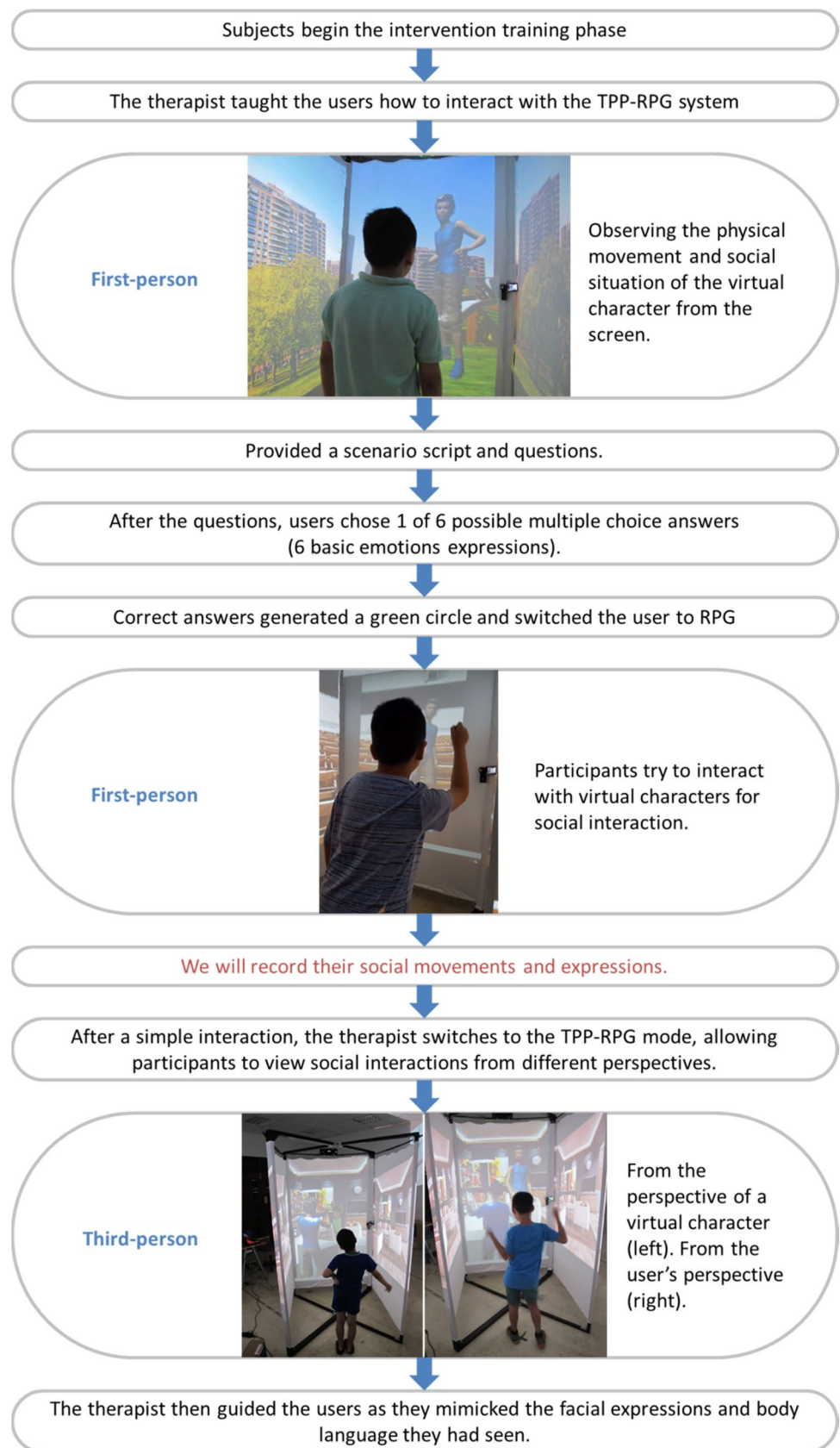


the screen. From the back, the children could try to figure out the emotion what was being expressed through the body language that they could observe.

4.3 Maintenance phase

Between the intervention and maintenance phases, there was a 6-week hiatus to reduce recall interference. This time delay

Fig. 4 Operating the TPP-RPG system



allowed us to compare the children's responses when they were tested again with their responses in the baseline phase to see whether they had maintained the skills that they had learned from the intervention phase.

5 Data collection and test reliability and validity

Data collection was done using the same methods and standards as required of all scientific experiments. The researcher who examined the procedural reliability of this study was the same certified occupational therapist who conducted all the tests. We followed related experimental methods used in other studies to train and test the children's ability to identify correct greeting behaviors [17]. The test procedure was designed to follow standard operating procedures for a therapist to ensure consistency in the processes and related controls (3D animation content, time, test questions, gesture completion, facial expressions, case criteria, and test environment). The same TPP-RPG strategy and context design were used to control the consistency of each story event to ensure that there were no unclear or emotionally confusing parts.

6 Results

The purpose of this study was to find out whether our RPG system with its TPP framework helps autistic children to observe their own and other people's emotions, expressions, and body movements. All outcomes were measured in two steps. The first step focused on the user's ability to respond appropriately to a set of six basic emotions by accurately identifying these expressions and the body movements associated with them. The second step was to translate these responses into performance statistics that enabled us to evaluate the data and provided us with feedback regarding all scores and body movements. The response assessment was read only after the video recording procedure had been completed. From the resulting table of scores for each aspect being evaluated, we averaged the results to get more

reliable and stable measures. The system utilized a set of pretests taken by ordinary/normal students of the same age (7–9 years old) as our autistic patients to determine whether the system's calibration was sound and to make sure our system was appropriate for the age and abilities of our subjects. In addition, expressions and body movements were recorded and presented to the therapist for reference to see whether the behavior of the child with autism appeared to have improved as a result of using the system.

6.1 Learning effects of the TPP-RPG training system

In each session of the experiment, 20 SSTs were given in three phases, and the scores of correct responses were recorded. Each training session had 20 SSTs. For each correct answer, participants received 2 points, for a total of 40 points. An occupational therapist and several researchers checked the answers and tested for normative answers. After the children had completed each test in each phase, we used questionnaires and interviews for expert assessment, and parental review related to the results of the tests to ensure their social reliability and that the tests were consistently measuring actual performance. In order to ensure that the results of the tests were reliable, experts in ASD and parents were asked to review the results of interviews and questionnaires and assess whether the tests accurately measured the children's behavior during the test when compared to their behavior in the real world.

Our system utilized a multiple-baseline design across single subjects [38]. The baseline phase consisted of four sessions for Doran, six for Bob, and four for Peter. The intervention phase consisted of 10 sessions for each child. The maintenance phase involved six sessions for Doran, four for Bob, and six for Peter. According to the single-case research method, the three children were required to perform 20 SSTs in each session.

At the beginning of the experiment, all three children started with low scores (mean score 11–23) during the baseline phase (Table 2). All participants' scores rose and continue to steadily improve (mean score 22–24) during the intervention phase. In the maintenance phase, two subjects (Bob and Peter) maintained scores higher than those at

Table 2 Correct response scores and percentage for participants

Participants	Performances of correct response scores and percentage					
	Baseline		Intervention		Maintenance	
Doran	11	27.5%	22	55.0%	11	27.5%
Bob	23	57.5%	24	60.0%	34	85.0%
Peter	13	32.5%	24	60.0%	30	75.0%
Mean	15.7	39.2%	23.3	58.3%	25.0	62.5%

In each training session there are 20 SSTs. If participants answer back one correct answer, they can get the 2 points, for a total of 40 points

their baselines (mean score 11–34). We found Doran’s score decreased in the maintenance phase because his learning status was less stable than the others. (The therapist indicated that when Doran was subjected to the traditional training strategy during the maintenance phase, he felt bored and was not willing to participate, so his scores were low.) During the intervention phase, however, his scores almost doubled, rising from 11 to 22, which was similar to the other two participants.

Overall, the data outcomes indicate that the mean correct assessment rates for the learning curve had improved after training and that in the maintenance phase the subjects retained most of the social expression and social skills that they had learned in the intervention phase.

6.2 Overall role-play performance

The training effect was independently evaluated by a therapist, who used a 5-point Likert scale (Table 3). The means of the scores for each phase were recorded, and it could be seen that all three children started with low scores (mean range 1.35–1.43) during the baseline phase. The independent therapist reported that all three scores rose significantly ($p < 0.05$) and dramatically (they more than doubled; mean range 3.76–3.96) during the intervention phase, and they remained significantly higher in the maintenance phase than in the subjects’ baseline assessment (mean range 3.47–3.93).

7 Discussion

7.1 Feedback from the children, therapist, and parents

During the experiment, we found that the children were curious and excited about their own images being shown on the projection wall, especially when they found that their actions could affect the virtual characters in the TPP-RPG training system. This type of interaction capability made them more interested in social reciprocity. They would often start exploring their social interactions with others on their own and ask the therapist “Am I doing this right?.” Because they could freely switch to see a perspective on

their behavior from the viewpoints of different people, they could repeat actions and explore them from different angles, which prompted them to think about the nature and importance of these socially reciprocal actions. This capability also increased the opportunities for them to practice and understand and helped them try to figure out the dynamics of basic social actions. Our TPP-RPG training system also appeared to help these ASD children empathize with others, understand body language better, and become observers of the behavior of other people with whom they were interacting.

After the intervention, the children continuously watched the 3D animations and focused on the virtual roles being presented to them. They mimicked the actions of the virtual characters. Although the children were unable to always mimic the exact facial expressions and body language that they observed during the first intervention, they tried to. The therapists and parents said that when they were faced with social encounters in their regular lives the children always shied away from interacting, or else they started looking for someone familiar to lean on. When the children were in a hyperactive state, they sometimes ignored or even run away from a classmate, neighbor, or stranger who waved hello. After some initial interventions using our system, the children began to look at facial expressions and body language and to ask their therapists what these meant. They then began to mimic them. The TPP-RPG system made learning interesting, fun, and productive for the children and helped them focus on greeting behaviors. We also found that their role-play skills were better after the interventions.

7.2 Benefits of using the TPP-RPG training system

In the present study, RPG, expressions of whole-body non-colloquial socialization, and body movements were all combined to train children with ASD. Using our TPP-RPG system, children with ASD could interact with virtual characters, and by overlapping their own characters and virtual roles from a TPP-enabled third-person perspective, they were also able to see their own facial expressions and body language in addition to that of the other characters. This capability allowed the users to think about the relationships between their own characters and those of others. The advantages of a TPP-RPG system lie in its integration of interaction and immersion, which cannot occur when participants in role-play are only looking at 2D pictures.

Additionally, we found that when our TPP-RPG system is used to train children with ASD to see, understand, and mimic facial expressions and body language it is measurably efficacious. Previous RPG training strategies primarily asked users to interact with the therapist or their teacher, but in such a straightforward context there is no way for users to watch the interaction between themselves and those with

Table 3 Role-play performance for participants

Participants	Performances of role-play skills		
	Baseline	Intervention	Maintenance
Doran	1.35	3.76	3.47
Bob	1.43	3.96	3.65
Peter	1.30	3.86	3.93
Mean	1.36	3.86	3.68

whom they are interacting. For children with ASD to see their own facial expressions and body language in 3D at the same time as they can observe the behavior of corresponding virtual characters is materially helpful. Indeed, it increases their motivation to learn and generates empathy.

7.3 Limitations and future work

Our investigation has at least four limitations. One is that a CAVE-like immersive 3D virtual reality environment is required, and therefore, a large-scale screen for projection must be erected, together with an aluminum frame and three cameras, which makes it difficult to set up and carry out in schools and to use with children with ASD. Even so, based on the speed with which current technology is evolving, smaller, less expensive, and more powerful projectors will become available, and the quality of the pictures they project can increasingly be improved. It is also conceivable that lighter and stronger materials will become available, making the transportation of the apparatus much easier.

Another limitation in our study is that the use of overlapping pictures of autistic children interacting with virtual characters presented from the third-person perspective can lead to confusion about what perspectives are involved. A side recording device is needed to assist autistic children by providing a history for them to watch after the training, to clear up any confusion involving sound–image variations from the real-time presentation.

A third limitation is that during systemic training, body movements can be better identified than facial expressions. However, both are equally important, but in this training case, our participants rarely had opportunities to focus on the facial parts due to our training system design. Almost by its fundamental nature, our perceptual system favors observations of spatial body interaction as opposed to the much more subtle aspects of facial expression. More discriminating training methods and materials are needed that can better focus on the highly nuanced aspects of what facial expressions may be appropriate in any given social situation. In future studies, it would be valuable to introduce new design elements that help subjects focus equally on facial expression. In this way, subjects could give equal weight to their observations of full body language and of facial expressions when attempting to identify the emotions that are being communicated.

Finally, our assessments are addressed at determining the improvement of a patient's overall status. However, it remains important to focus in much greater detail on the many differences between individuals, including the influence of gender, age, initial IQ scores, and social behavior.

Such future work could include case studies of individual emotions and their impact on how children respond to our training system. This would provide significantly more subtle and detailed data for analysis.

8 Conclusion

The study presented in this paper used a TPP-RPG strategy to develop an immersive and interactive training system aimed at teaching children with ASD how to observe, understand, and appropriately react to the facial expressions and body language of their interlocutors. The proposed system allows its users to observe their own and other characters' facial expressions and body language. This is the first study regarding the use of a CAVE-like immersive 3D virtual reality environment to provide ASD children with a third-person perspective. Overall, we found that the TPP-RPG approach developed was measurably and significantly effective in achieving the aims specified. The CAVE-like immersive environment provides a very good training method and is more efficient than the graphics card recognition system which is generally used. In addition, compared with complex VM systems, the technology and materials used now are easier to obtain and the learning efficiency which they support is also better.

In particular, the system led to observable and moderate changes in the ability of the three children involved to recognize and understand the facial expressions and body language of the real and virtual interlocutors with which they were involved in the study. It also made learning how to understand social situations, social cues, and social behavior interesting and enjoyable and was clearly more effective compared to traditional methods. Our intervention system was also productive in helping the children to maintain their focus and therefore better recognize affective expressions, and to generally promote their social skills. Using our system triggered their incentive to learn and encouraged them to observe nonverbal social and emotional signals and to improve their role-play skills. In future studies, experiments might want to use our TPP-RPG system and include non-ASD participants of all ages together with ASD participants to spur research in this area. In addition, training materials for ASD individuals need to be more complete and more reflective of real life, and we suggest that our system moves us forward in this regard. Finally, we hope that our findings will provide guidance to new research projects on how to create visual media that can increase the ability of adolescents and others with ASD to recognize nonverbal social reciprocity cues in a wide variety of interpersonal situations.

Acknowledgements We are grateful to the participants, therapists, and family members who participated in the study as well as the participants who assisted in the various phases of the study. We would also like to thank the individuals who participated in this research and the Autism and Developmental Research Center in Taiwan.

Funding This work is supported by the Ministry of Science and Technology of Taiwan (MOST 107-2218-E-027 -013 -MY2). The authors thank the referees very much for their valuable comments and suggestions on this paper.

Appendix 1

Parts of the social scripts used in the experiments

	Scenes	Related character	Social scripts	Related character feelings	Respond to each other’s social actions
1	Park	Friends	Joe borrowed a toy car from Tom. Joe accidentally broke a wheel while playing with the car. Tom saw it and said: “Hum! You broke my toy car. I don’t want to be nice to you!”	Anger	Apologize (bow)
2	Park	Friends	When Tina was playing in the park, she met her new friend Allen. Allen said to Tina: “Hi, we will play together in the future, we will be good friends.”	Happiness	Hold hands
3	Living room	Mother	Peter was at home because his mother had gone shopping. She was gone a long time and he had to spend a long time waiting for her. “Why didn’t my mother come back yet? Why is she so late?”	Fear	Ask hugs
4	Living room	Mother	Today, my mother cooked John’s favorite food for dinner. John said at dinner: “Awesome, I am getting my favorite—half-boiled eggs tonight!”	Happiness	Ask hugs

	Scenes	Related character	Social scripts	Related character feelings	Respond to each other’s social actions
5	Park	Friends	On the way home, when Lion was going through an alley, he saw his good friend Ali suddenly appears. “Wow! Why are you hiding here? You scared me!” he exclaimed.	Surprise	Express surprise (action)
6	Living room	Mother	When Lina came home from practicing riding a bicycle, her mother saw her palm was injured. Her mother asked what happened to Lina and she said: “I accidentally fell and it hurts!”	Sadness	Ask hugs
7	Class room	Friends	Gino and Archie are good friends and both like animations. At school, Gino said to Archie: “Did you watch that animation yesterday? Super funny!”	Happiness	High five
8	Class room	Friends	Bob borrowed a pencil with Joanna, but Bob lost the pencil. Joanna complained to the teacher: “Teacher, he lost my pencil!”	Disgust	Raise hand
9	Park	Friends	Danny was lost in a hide-and-seek game with friends and said: “Where am I? I have accidentally run into a strange place”	Fear	Ask for help
10	Class room	Classmate	Bill made an appointment with his classmates to play after class, but when he was about to go out he saw that it was raining: “Oh! Why is it raining ~ Now I can’t go out to play!”	Sadness	Talk to classmate (action)

References

- Adjorlu, A., Hussain, A., Mødekjær, C., Austad, N.W.: Head-mounted display-based virtual reality social story as a tool to teach social skills to children diagnosed with autism spectrum disorder. In: 2017 IEEE Virtual Reality Workshop on K-12 Embodied Learning Through Virtual & Augmented Reality (KELVAR). IEEE (2018)
- Ahlers, K.P., Gabrielsen, T.P., Lewis, D., Brady, A.M., Litchford, A.: Supporting individuals with autism spectrum disorder in understanding and coping with complex social emotional issues. *Sch. Psychol. Int.* **38**(6), 586–607 (2017)
- American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). American Psychiatric Pub, Philadelphia (2013)
- Anderson, P.L., Rothbaum, B.O., Hodges, L.: Virtual reality: using the virtual world to improve quality of life in the real world. *Bull. Menn. Clin.* **65**(1: Special issue), 78–91 (2001)
- Axe, J.B., Evans, C.J.: Using video modeling to teach children with PDD-NOS to respond to facial expressions. *Res. Autism Spect. Disord.* **6**(3), 1176–1185 (2012)
- Ayres, K.M., Langone, J.: Intervention and instruction with video for students with autism: a review of the literature. *Educ. Train. Dev. Disabil.* **40**, 183–196 (2005)
- Baron-Cohen, S.E., Tager-Flusberg, H.E., Cohen, D.J.: *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience*. Oxford University Press, Oxford (2000)
- Baron-Cohen, S., Belmonte, M.K.: Autism: a window onto the development of the social and the analytic brain. *Annu. Rev. Neurosci.* **28**, 109–126 (2005)
- Baron-Cohen, S., Leslie, A.M., Frith, U.: Does the autistic child have a “theory of mind”. *Cognition* **21**(1), 37–46 (1985)
- Baker, J.: *The Social Skills Picture Book: Teaching Communication, Play and Emotion*. Future Horizons, Arlington (2017)
- Becker-Weidman, A.: Effects of early maltreatment on development: a descriptive study using the Vineland Adaptive Behavior Scales-II. *Child Welf.* **88**(2), 137–161 (2009)
- Bellini, S., Akullian, J.: A meta-analysis of video modeling and video self-modeling interventions for children and adolescents with autism spectrum disorders. *Except. Child.* **73**(3), 264–287 (2007)
- Bierman, K.L., Greenberg, M.T., Coie, J.D., Dodge, K.A., Lochman, J.E., McMahon, R.J.: *Social and Emotional Skills Training for Children: The Fast Track Friendship Group Manual*. Guilford Publications, New York (2017)
- Bölte, S., Hallmayer, J. (eds.): *Autism Spectrum Conditions: FAQs on Autism, Asperger Syndrome, and Atypical Autism Answered by International Experts*. Hogrefe Publishing, Göttingen (2011)
- Cameirão, M.S., i Badia, S.B., Oller, E.D., Verschure, P.F.: Neurorehabilitation using the virtual reality based rehabilitation gaming system: methodology, design, psychometrics, usability and validation. *J. Neuroeng. Rehabil.* **7**(1), 48 (2010)
- Charlop-Christy, M.H., Daneshvar, S.: Using video modeling to teach perspective taking to children with autism. *J. Posit. Behav. Interv.* **5**(1), 12–21 (2003)
- Charlop, M.H., Lang, R., Rispoli, M.: Lights, camera, action! Teaching play and social skills to children with autism spectrum disorder through video modeling. In: Charlop, M.H., Lang, R., Rispoli, M. (eds.) *Play and Social Skills for Children with Autism Spectrum disorder*, pp. 71–94. Springer, Cham (2018)
- Chen, C.H., Lee, I.J., Lin, L.Y.: Augmented reality-based self-facial modeling to promote the emotional expression and social skills of adolescents with autism spectrum disorders. *Res. Dev. Disabil.* **36**, 396–403 (2015)
- Chen, C.H., Lee, I.J., Lin, L.Y.: Augmented reality-based video-modeling storybook of nonverbal facial cues for children with autism spectrum disorder to improve their perceptions and judgments of facial expressions and emotions. *Comput. Hum. Behav.* **55**, 477–485 (2016)
- Cohn, J.F., Ambadar, Z., Ekman, P.: Observer-based measurement of facial expression with the facial action coding system. In: Coan, J.A., Allen, J.J.B. (eds.) *The Handbook of Emotion Elicitation and Assessment*, vol. 1, no. 3, pp. 203–221. Oxford University Press, New York (2007)
- Davies, R.R., Rogers, E.S.: Social skills training with persons who are mentally retarded. *Ment. Retard.* **23**(4), 186 (1985)
- Ekman, P.: Emotional and conversational nonverbal signals. In: Larrazabal, J.M., Miranda, L.A.P. (eds.) *Language, Knowledge, and Representation*, pp. 39–50. Springer, Dordrecht (2004)
- Ghanouni, P., Jarus, T., Zwicker, J.G., Lucyshyn, J., Mow, K., Ledingham, A.: Social stories for children with autism spectrum disorder: validating the content of a virtual reality program. *J. Autism Dev. Disord.* **49**(2), 660–668 (2019)
- Goldstein, S., Ozonoff, S. (eds.): *Assessment of Autism Spectrum Disorder*. Guilford Publications, New York (2018)
- Gray, C.: *Social Stories*. Future Horizons, Arlington, TX (1994)
- Hillier, A., Fish, T., Cloppert, P., Beversdorf, D.Q.: Outcomes of a social and vocational skills support group for adolescents and young adults on the autism spectrum. *Focus Autism Other Dev. Disabil.* **22**(2), 107–115 (2007)
- Howlin, P., Yates, P.: The potential effectiveness of social skills groups for adults with autism. *Autism* **3**(3), 299–307 (1999)
- Ip, H.H., Wong, S.W., Chan, D.F., Byrne, J., Li, C., Yuan, V.S., Lau, K.S.Y., Wong, J.Y.: Enhance emotional and social adaptation skills for children with autism spectrum disorder: a virtual reality enabled approach. *Comput. Educ.* **117**, 1–15 (2018)
- Karal, M.A., Wolfe, P.S.: Social story effectiveness on social interaction for students with autism: a review of the literature. *Educ. Train. Autism Dev. Disabil.* **53**(1), 44–58 (2018)
- Lee, S.Y., Lo, Y.Y., Lo, Y.: Teaching functional play skills to a young child with autism spectrum disorder through video self-modeling. *J. Autism Dev. Disord.* **47**(8), 2295–2306 (2017)
- Lee, I.J., Chen, C.H., Wang, C.P., Chung, C.H.: Augmented reality plus concept map technique to teach children with ASD to use social cues when meeting and greeting. *Asia Pac. Educ. Res.* **27**(3), 227–243 (2018)
- Lee, I.J., Lin, L.Y., Chen, C.H., Chung, C.H.: How to create suitable augmented reality application to teach social skills for children with ASD. In: Mohamudally, N. (eds.) *State of the Art Virtual Reality and Augmented Reality Knowhow*, vol. 8, pp. 119–138. IntechOpen, London (2018)
- Lillard, A.S.: Why do the children (pretend) play? *Trends Cogn. Sci.* **21**(11), 826–834 (2017)
- Nikopoulos, C.K., Keenan, M.: Using video modeling to teach complex social sequences to children with autism. *J. Autism Dev. Disord.* **37**(4), 678–693 (2007)
- Parsons, S., Leonard, A., Mitchell, P.: Virtual environments for social skills training: comments from two adolescents with autistic spectrum disorder. *Comput. Educ.* **47**(2), 186–206 (2006)
- Pearson, A., Ropar, D., Hamilton, A.F.D.: A review of visual perspective taking in autism spectrum disorder. *Front. Hum. Neurosci.* **7**, 652 (2013)
- Reynhout, G., Carter, M.: Social Stories™ for children with disabilities. *J. Autism Dev. Disord.* **36**(4), 445–469 (2006)
- Richards, S.B.: *Single Subject Research: Applications in Educational Settings*. Cengage Learning, Boston (2018)
- Riva, G.: Virtual reality in psychotherapy. *Cyberpsychol. Behav.* **8**(3), 220–230 (2005)

40. Sharples, S., Cobb, S., Moody, A., Wilson, J.R.: Virtual reality induced symptoms and effects (VRISE): comparison of head mounted display (HMD), desktop and projection display systems. *Displays* **29**(2), 58–69 (2008)
41. Shillingsburg, M.A., Cariveau, T.: Responding to conversational partners' behavior: teaching children with autism strategies to recapture the interest of others. *Evid. Based Commun. Assess. Interv.* **10**(3–4), 143–147 (2016)
42. Smyth, J.M.: Beyond self-selection in video game play: an experimental examination of the consequences of massively multiplayer online role-playing game play. *CyberPsychol. Behav.* **10**(5), 717–721 (2007)
43. Tustin, F.: *Autistic Barriers in Neurotic Patients*. Routledge, London (2018)
44. Van Wermeskerken, M., Grimmius, B., van Gog, T.: Attention to the model's face when learning from video modeling examples in adolescents with and without autism spectrum disorder. *J. Comput. Assist. Learn.* **34**(1), 32–41 (2018)
45. Wallace, S., Parsons, S., Westbury, A., White, K., White, K., Bailey, A.: Sense of presence and atypical social judgments in immersive virtual environments: responses of adolescents with autism spectrum disorders. *Autism* **14**(3), 199–213 (2010)
46. Yuan, S.N.V., Ip, H.H.S.: Using virtual reality to train emotional and social skills in children with autism spectrum disorder. *Lond. J. Prim. Care* **10**(4), 110–112 (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.