Enhancing Social Skills of children with HFA Using Assistive Technology

Maha Abdelmohsen

A thesis submitted in partial fulfilment of the requirements of the University of Greenwich for the Degree of Doctor of Philosophy

August 2021

DECLARATION

I certify that the work contained in this thesis, or any part of it, has not been accepted in substance for any previous degree awarded to me or any other person, and is not concurrently being submitted for any other degree other than that of Doctor of Philosophy which has been studied at the University of Greenwich, London, UK. I also declare that the work contained in this thesis is the result of my own investigations, except where otherwise identified and acknowledged by references. I further declare that no aspects of the contents of this thesis are the outcome of any form of research misconduct.

I declare any personal, sensitive or confidential information/data has been removed or participants have been anonymised. I further declare that where any questionnaires, survey answers or other qualitative responses of participants are recorded/included in the appendices, all personal information has been removed or anonymised. Where University forms (such as those from the Research Ethics Committee) have been included in appendices, all handwritten/scanned signatures have been removed.

Student Name: Maha Abdelmohsen
Student's Signature:

Date: 23/08/2021

First Supervisor's Name: Yasmine Arafa
First Supervisor's Signature:

Date: 23/08/2021

Second Supervisor's Name: Cornelia Boldyreff
Second Supervisor's Signature:

Date: 23/08/2021

ACKNOWLEDGEMENTS

I would like to acknowledge everyone who played a role in my PhD journey. First of all, I would like to thank my supervisors, Dr. Yasmine Arafa, Prof. Cornelia Boldyreff, and Prof. Jixin Ma, for their great support, motivation, and guidance through this journey. Thanks for all the inspiring ideas and constructive comments. The completion of this study could not have been possible without your knowledge and experience. Secondly, I would like to express my gratitude to my supportive friends, who supported and encouraged me all the time. I would like to thank my mother for the unconditional love and endless support I witnessed during my PhD journey; without you I could never have reached any success. Finally, I would like to express my gratitude and appreciation to all the parents who participated in this research with their children.

Abstract

Social skills training (SST) interventions are crucial to train and enhance the social skills of children with autism spectrum disorder (ASD). SST interventions provide learning experiences to teach children the necessary skills to interact successfully with their social environment. Although traditional SST interventions were proven to be beneficial for children with ASD by many previous studies, their availability is limited. In recent years, physical social robots and virtual environments (VE) have been popular training tools for children with ASD. Studies were conducted to evaluate their effectiveness in enhancing the social skills of children with ASD. This research investigates the impact of combining a virtual environment with a virtual social robot (VRobot) as a hybrid approach to train the social skills of children with high-functioning autism (HFA). A non-immersive virtual environment (Jammo-VRobot environment) that employs a 3D robot (Jammo VRobot) was designed. Jammo-VRobot environment aims to enhance the social skills of children with HFA through a social skills training programme guided by a parent or a teacher. This research aims to provide a tool that can be widely available, cheap, and easily used by parents and teachers either at home or school. The designed social skill training programme is adapted and modified from successful work done with physical robots (NAO, Zeno, and Qtrobot) to train the social skills of children with ASD. The developed training programme targets three social skills: imitation, emotion recognition and expression, and intransitive gestures. The evaluation process was conducted mostly online with some on-site, including 15 children with HFA (aged 4-12 years). The participants were taught to recognise six basic emotions and 11 intransitive gestures (Phase I), to imitate these emotions and gestures (Phase II), and to express and demonstrate them in social situations (Phase III). The experimental sessions reveal encouraging results showing that the Jammo-VRobot environment helps in training and enhancing the target three skills of the participants. Additionally, similar results were found by comparing the results achieved through interaction with Jammo VRobot and those achieved through interaction with physical robots used by SoA. The findings indicate that Jammo-VRobot is as effective as the physical robots in training and enhancing the social skills of children with HFA. It also emphasises that the Jammo-VRobot environment overcame some limitations of social robots and immersive VE environments in terms of availability and cost.

Contents

List of Tables

List of Figures

1 Introduction

1.1 Motivation

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that appears at an early age and affects child development (Kanner et al., 1943; Rao et al., 2008). While ASD is a spectrum disorder, characteristics can vary from one child to another. Each child with ASD may have different characteristics or a different combination of characteristics. However, some difficulties are commonly observed in them. These difficulties include communication impairments, repetitive behaviours, little or delay language skills, and social impairment. Social skills are fundamental for maintaining the quality of life and interaction with others. Impairments in social skills are a cardinal characteristic of ASD. Children with ASD are classified according to the severity of their difficulties, such as Asperger's Syndrome, Autistic Disorder, and Pervasive developmental disorder. Asperger Syndrome is known as high-functioning ASD with mild characteristics. High-functioning autism (HFA) refers to children who have been diagnosed with ASD with an Intelligence Quotient (IQ) above 70 (Attwood, 2003; Malinverni et al., 2017). Although children with HFA exhibit higher language development skills, they may have difficulties in social skills (Society, 2016; Attwood, 2003). The autistic disorder is also known as low-functioning ASD (LFA) with significant language delays and below the average intelligence. Pervasive developmental disorder combines some characteristics between HFA and LFA with fewer and milder difficulties (Mandan, 2019). HFA and LFA are not medical diagnoses and not formal subcategories of autism spectrum disorder (Estratopoulou and Sofologi, 2019). Therefore, this research focuses on enhancing the social skills of children with high-functioning ASD.

Success in improving the social skills of children with ASD may be achieved through early interventions. Additionally, it is easier to teach the appropriate behaviour and new skills to young children when the brain is most easily developed (So et al., 2018b). Thus, the target age group for this research is 4-12 years old.

There is no cure for ASD (Happé et al., 2006; Dapretto et al., 2006), although studies report that seven months of traditional social skill training sessions have improved the social and communication skills of children with ASD (Rao et al., 2008). Therapists or specialist teachers facilitate the traditional social skill training sessions. Picture Exchange Communication System (PECS), Applied Behaviour Analysis (ABA), Occupational Therapy (OT), social skills groups, social stories, video-modelling, and peer-mediated interventions are traditional behavioural interventions (Bohlander et al., 2012; Makrygianni et al., 2018). However, these therapeutic interventions are costly and are not always accessible because of the limited number of providers (Rosenfield et al., 2019). With the rapid growth of technology, a wide range of studies has investigated the potential use of assistive technology as an alternative means for children with ASD to acquire social skills. Assistive technology is any device, equipment, software, or product system that is used to increase, maintain, or improve the functional capabilities of person with disabilities (Assistivetechnology, 2022). This research focuses only on the use of assistive technology for individuals with cognitive disabilities. Assistive technology used in training the social skills of children with ASD is known as an assistive technology for cognition (ATC). ATC has been widely used to support individuals with cognitive impairments. Virtual Environments (VEs) and social robots are types of assistive technology for cognition that have been used in training the social skills of children with ASD.

Virtual environments have become a widespread tool for ASD training in recent years. Studies were conducted to evaluate the effectiveness of virtual environments in training the social skills of children with ASD (Nojavanasghari et al., 2017; Bekele et al., 2013; Serret et al., 2014). Virtual environments are used to train the social skills of children with ASD are classified into immersive virtual environments and non-immersive (desktop) virtual environments. In the immersive virtual environment, the user is completely surrounded by 3D objects. In the non-immersive virtual environment, the user interacts with digital content through displays without being immersed. Studies show that VE is a promising tool for enhancing the social skills of children with ASD. However, there is no clear evidence that participants can transfer the learnt skills in the VE to the real world (Shahab et al., 2017), as these studies have not conducted follow-up sessions to assess whether the children transferred the gained skills to their daily life. The lack of younger participants in these studies is due to the lack of autistic schools and the long process of obtaining parents' approval through completing consent forms. Additionally, the high cost of immersive virtual environment equipment is another limitation of the widespread of immersive VE interventions.

Studies using robot-based interventions with children with ASD report their positive effect on the child's engagement (Huijnen et al., 2017). Using robots can enhance the learning outcome, as engagement is considered as an essential pre-requisite for learning, where higher engagement results in more opportunities for social skill learning (Schadenberg et al., 2020; Costa et al., 2017; Huijnen et al., 2017; Zorcec et al., 2018). Despite the beneficial outcomes of social robot interventions, not all families with a child with ASD can afford the robot-assisted sessions, as these interventions are costly. Furthermore, the set-up of this environment requires a professional or a technician. Additionally, not all countries have robot-assisted interventions.

Successes achieved by using VEs and social robots interventions separately led to the idea that combining both techniques may address some limitations of these interventions and enhance the social skills of children with ASD. Thus, this PhD aims to develop an environment (Jammo-VRobot environment) that combines both techniques and investigates the impact of a desktop (non-immersive) VE that utilises a virtual robot (Jammo VRobot) in enhancing the social skills of children with HFA. The intention is to produce a promising and cost-effective platform that parents and teachers can easily use to practise social scenarios for HFA.

1.2 Research Question

The following research questions are aimed to be explored in this research:

- What are the alternative methods for improving the social skills of children with ASD?
- How can a virtual environment (VE) that combines two technologies (VE and Robot) contribute to improving the social skills of children with high-functioning ASD?
  - Could the virtual robot be as effective as the physical robots used by SoA in enhancing the social skills of children with ASD?
  - To what extent does the Jammo-VRobot environment help children with HFA recognise and comprehend social emotions, improve their imitation skills, and recognise and produce intransitive gestures?
  - Can the Jammo-VRobot environment help children with HFA generalise the skills they learnt in the intervention sessions to their daily life?

1.3 Objectives

- Combine two assistive technologies, VE and social robots, for social skill training.
- Use these technologies in association with common social scenarios for developing a social skill training tool aimed at children with HFA.
- Develop a tool that parents and teachers can use and control without interference from a researcher.
- Evaluate the Jammo-VRobot tool with participants (4-12 years) on-site and online to investigate the approach's efficacy.
- Analyse the data collected from the experimental sessions to evaluate the effectiveness of the Jammo-VRobot environment.

1.4 Hypotheses

The robot-assisted intervention has gained traction in the past few years to become an emerging and promising application area in the education of children with ASD and has shown promising results in training their social skills (So et al., 2018b; Costa et al., 2017; Marino et al., 2020). Using virtual environments in training children with ASD has been a focus of research over the past two decades. This form of technology provides a safe, controllable, and repeatable environment for children with ASD to practise their social skills (Lorenzo et al., 2019; Stewart Rosenfield et al., 2019; Bozgeyikli et al., 2018). Developing a virtual environment that combines VE and a social robot that utilises established social skill training programmes could effectively enhance their social skill. Providing an open-source tool that is cheap, widely available, and easily used by parents and teachers could overcome the geographically constrained availability of social robots and immersive VE.

Testing generalisation of the skills is crucial to evaluate the effectiveness of any social skill training intervention. Generalisation is transferring the learnt skills in specific environments to a novel (new) environment. Involving parents and teachers in the intervention sessions and providing follow-up sessions could allow the researcher to assess the generalisation of the learnt skills to real-life (Khowaja et al., 2020). Parents and teachers are involved in the daily life of the children, which helps to give better insight into the effectiveness and the impact of the Jammo-VRobot environment in enhancing the social skills of children with HFA. Additionally, they could notice if the children applied what they have learnt in their daily life.

- It is hypothesised that the combination of the two assistive technologies, VE and Social Robot, will be as effective as the real (physical) robot in enhancing the social skills of children with HFA, such as imitation skills, intransitive gestures, and train their emotion ability.
- Children with HFA can generalise the learnt skills, imitation, emotion recognition, recognition, and production of intransitive gestures within the Jammo-VRobot environment to other social contexts.
- The Jammo-VRobot environment could overcome the limitation of social robots in terms of cost and availability and become a widely available tool and easy to use by parents and teachers.

1.5 Contribution

This study aims to evaluate the effectiveness of the Jammo-VRobot environment to enhance the social skills of children with high-functioning ASD and provide a widespread tool. The contributions are:

- A novel approach to enhance the social skills of children with high-functioning ASD through the combination of two assistive technologies: Social Robot and Virtual Environment, with an established social skills training programme.
- Development of an open-source VE that is widely available and can be easily used by parents and teachers to practise scenarios that aim to enhance the social skills of children with high-functioning ASD.
- Comparison between the results achieved through interactions with the virtual robot (Jammo) and those achieved through interaction with physical robots used by the state-of-the-art.

1.6 Developing of the Research

To answer the research questions, extensive research on autism, deficits in autism, social skills training programmes, and assistive technologies were surveyed and critically analysed. The initial investigation began with understanding the definition of autism spectrum disorder and its deficits. Social communication skills are a core feature of children with ASD. Therefore, the traditional social skills training interventions that have been frequently used in training the social skills of children with ASD were investigated. Although these traditional interventions have achieved promising results, their availability is limited.

Animal-assisted interventions were investigated. Children with ASD seem more comfortable, socially motivated, and engage more naturally while interacting with dogs than humans. However, animal-assisted interventions are limited due to the high cost of training animals, safety, and ethical concerns.

The literature review provided clear evidence that assistive technology is becoming a promising solution to overcome the limitations of traditional training programmes and animal-assisted interventions. After an in-depth investigation into the use of assistive technologies in enhancing the social skills of children with ASD, virtual environments (VE) and social robots became the focus of this research.

The research into the use of virtual environments and social robots in enhancing the social skills of children with ASD emphasised some limitations and led to the proposed approach, which combines a virtual environment with a social robot to address some of their limitations.

Due to the positive effect of the animal-assisted intervention, especially dogs, on children with ASD, the intention was initially to develop a 3D robotic dog that can interact with the child. The MiRo virtual dog that was developed by the University of Sheffield was considered. However, the limited number of its animation and the limited number of scenarios that can be applied stated a problem. Thus, the decision was made to use a 3D humanoid robot.

The design of the social scenarios was the next step towards the development of the environment. Social skill interventions often involve training imitation, turn-taking, emotion recognition and expression, joint attention, and gestures. Imitation, emotion recognition and expression, and intransitive gestures were chosen as the research's target skills. The deficits of these skills are associated with the deficits of the social skills of children with ASD.

To evaluate the effectiveness of the Jammo-VRobot environment, recruiting participants was essential. However, recruiting children with HFA to participate in the study is challenging. Firstly, ethical approval was obtained from the University of Greenwich Research Ethics Committee (UREC). Specialist schools in Egypt were contacted for the evaluation process. Due to Covid-19 and the lockdown, the evaluation had to be conducted online as well as on-site. Thus, an online version of the Jammo-VRobot environment was launched on a website to be available for a wider group.

Data collection methods were reviewed to evaluate the effectiveness of the Jammo-VRobot environment in enhancing the social skills of children with HFA. A hybrid method of qualitative and quantitative data collection was used.

Data of the participant's interaction was gathered through observations and pre and post questionnaires. The data were analysed using the Wilcoxon statistical method.

1.7 Publications and Talks

- A workshop paper was submitted to "Virtual Reality and 3D User Interfaces, IEEE Conference" and got accepted and published in IEEE proceedings. The paper is titled "Training Social Skills of Children with ASD Through Social Virtual Robot".

  Abdelmohsen, M. and Arafa, Y., 2021, March. Training Social Skills of Children with ASD Through Social Virtual Robot. In 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) (pp. 314-319). IEEE.

- A conference paper was submitted to "ACM International Conference on Information Technology for Social Good (GoodIT 2021)" and got accepted and published in ACM proceedings. The paper is titled "Virtual Jammo Virtual Robot Enhances the Social Skills of Children with HFA: Development and Deployment."

  Abdelmohsen, M. and Arafa, Y., 2021, September. Jammo Virtual Robot Enhances the Social Skills of Children With HFA: Development and Deployment. In Proceedings of the Conference on Information Technology for Social Good (pp. 180-185).

- Another workshop paper got accepted in "Workshop at International Conference on Social Robotics (ICSR) 2020". The paper titled "Virtual Social Robot Interactions for Enhancing the Social Skills of Children with ASD". It is available on the workshop website.

  `https://sites.google.com/view/icsr-2020-workshop/accepted-papers`

- A conference paper was submitted to "The 13th International Conference on Social Robotics (ICSR 2021)", got accepted and published by Springer. The paper titled "Virtual Social Robot Enhances the Social Skills of Children with HFA".

  Abdelmohsen, M. and Arafa, Y., 2021, November. Virtual Social Robot Enhances the Social Skills of Children with HFA. In International Conference on Social Robotics (pp. 497-508). Springer, Cham.

- Finalising a journal paper to be submitted to "Computers in Human Behavior" by Elsevier.

- A presentation was given to BCS: open-source video games

Here is a link for the talk: https://www.youtube.com/watch?v=kDNEYFOnn3I

- Another presentation was given to BCS: open-source in student activities

  Here is a link for the talk: https://www.youtube.com/watch?v=EDu9L_HOBd8&list=PLxde5XJWZRbQzntGfVtzmZqvJOPeapZ8r&index=6

1.8 Thesis Overview

The rest of this thesis is organised as described below:

- Chapter 2 presents an overview of the autism spectrum disorder and its challenges: theory of mind, communication skills, emotional skills, and social skills. The social skills training interventions that are commonly used for children with ASD are presented.
- Chapter 3 presents a literature review of the assistive technologies used for training children with ASD. Virtual environments and social robots training applications for children with ASD are reviewed. An overview of the assessment methods for evaluating the social skills of children with ASD is included.
- Chapter 4 presents the followed design considerations for developing the Jammo-VRobot environment.
- Chapter 5 provides a detailed description of the social skill training programme used in the Jammo-VRobot environment.
- Chapter 6 provides a detailed description of the experimentation and evaluation, including a description of the participants, setting, procedure, data collection methods, and the study results.
- Chapter 7 presents a discussion of the findings, future work, and conclusion.

2 Overview of Autism, Its impairments and Social skills training interventions

2.1 Chapter Overview

This review begins with the definition of Autism Spectrum Disorder (ASD) and a list of the impairments the disorder causes in terms of the theory of mind, emotions, and communication and social skills. The social skill training (SST) interventions used for training children with ASD are then reviewed.

2.2 Autism Spectrum Disorder

Autism Spectrum Disorders (ASD) represent a spectrum of developmental disorders characterised by impairments of verbal and non-verbal behaviours (Davis, 2018). Much is still unknown about what exactly causes ASD. However, researchers indicate that genetic factors such as fragile X syndrome, PKU or tuberous sclerosis and neurobiological factors contribute to the ASD stereotypical neurocognitive functions that characterise children with ASD (Rogers et al., 2001; Demetriou et al., 2018). In 2018, the Autism and Developmental Disabilities Monitoring Network (ADDM), funded by the Centres for Disease Control and Preventions (CDC), determined that approximately 1 in 59 children in the United States and 1 in 100 children in the United Kingdom were diagnosed with ASD (CDC, 2020). The statistics show that boys are four times more likely to be diagnosed with ASD than girls. ASD comprise of at least three subgroups: Autistic disorder (known as low-functioning), Pervasive developmental disorder (Kanner et al., 1943; The National Autistic Society, 2016), and Asperger syndrome (known as high-functioning) (Asperger and Frith, 1991). Children with ASD are characterised by the lack of or delay in spoken language, repetitive behaviour, little or no eye contact, lack of interest in others, lack of joint attention, and impairments of verbal and non-verbal behaviour (e.g. gestures, facial expressions, body postures) (Masi et al., 2017; Rao et al., 2008; Chen et al., 2016). Atypical eye contact and atypical joint attention are symptoms of children with ASD, affecting their social lives. They also have difficulty recognising others' emotions and intentions and taking appropriate decisions based on these emotions (Lee et al., 2018). These deficits can lead to behavioural challenges, such as self-injury and aggressive behaviour. For example, head banging, hand biting, and excessive self-rubbing and scratching (Edelson, 2016). According to Lance et al. (2014), repetitive behaviours might cause tissue damage, discolouration, and scarring over time.

Although children with ASD have many difficulties, they also have significant strengths. They can have the right eye for detail, a well-organised work environment, excellent memory, creative talent, and a high level of accuracy and reliability (Williams and Gray, 2013).

2.2.1 Theory of Mind (ToM) in ASDs

"Theory of mind is the ability to understand mental states, beliefs, emotions, intents, and desires of your own and others" (Saracho, 2014). ToM helps predict others' future behaviour by understanding their beliefs and how they feel (Baron-Cohen, 2008). Children with ASD have difficulty understanding others' emotional states, which contributes to the deficits in their theory of mind (Frith, 2001).

ToM is also a contributing factor to empathy. Empathy is an affective skill that helps understand the surrounding world and build an everyday social life (Happé et al., 2006; Dapretto et al., 2006). Empathising is recognising others' feelings and sensitivity to the different beliefs and perspectives of others (Lawson et al., 2004) and responding with the relevant emotion (Baron-Cohen, 2008). One of the main psychological characteristics of ASD is the lack of empathy (Trimmer et al., 2017).

2.2.2 Communication skills in ASDs

ASD remains a behavioural condition that causes behaviour and communication problems. Children with ASD find it difficult to have direct interaction with others. Neuroanatomical observations have shown that the brain cells responsible for social communication engagement are also responsible for the non-verbal social signals such as recognition and expression of feelings, eye gaze, and gestures (Bauman and Kemper, 2005). The abnormalities in these brain cells handle most of the characteristics of individuals with ASD (Happé et al., 2006; Rice et al., 2015). At pre-school, children with ASD can be distinguished from their peers by their social communication difficulties. They start to have a significant problem at elementary school in building friendships with same-age peers (Rao et al., 2008).

2.2.3 Emotions and ASD

Emotion is a critical element in establishing effective communication with people. Emotional awareness is the ability to understand feelings and influences the way people communicate. Humans express emotions through actions, which can be perceived through the visual, auditory and tactile modalities (Bartneck, 2002). Children with ASD have difficulties in understanding other's emotions. "They process faces differently and show reduced attention to faces and facial expressions" (Habash, 2014). They focus more on mouths rather than eyes (Habash, 2014). According to Ekman (1999), there are six basic emotions: fear, anger, surprise, joy, disgust, and sadness. Another subset of emotions is complex emotions, which include jealousy, pride, and shame. According to Habash (2014), typically developing children can recognise simple and complex emotions between the age of two and seven. In contrast, children with ASD have difficulties recognising basic emotions and struggle with recognising complex emotions into adulthood. Impairments in facial emotion recognition (FER) and facial emotion expression (FEE) are contributing factors to the poor social skills and difficulties in initiating friendships, which are the fundamental challenges that face children with ASD (Manfredonia et al., 2019; Rice et al., 2015; Happé et al., 2006; Dapretto et al., 2006).

2.2.4 Social Skills in ASDs

Social skills are the skills that humans use to communicate and interact with each other, verbally and non-verbally, through eye contact, gestures, facial expressions, and body language. According to Gresham et al. (2001), social skills are the skills that need to be taught and learnt for children with ASD to engage and communicate with others. Social skills deficits are another core characteristic of children with ASD (Barry, 2015; Gresham et al., 2001; Happé et al., 2006). These deficits affect their learning and playing skills. Children with ASD have difficulties observing social cues, stimuli, interaction, and standard warning, which may make them more vulnerable to danger.

The impairments in social skills of children with ASD classify into four categories: (a) emotion skills, which include the inability to recognise other's emotions from facial expression and body language; (b) play skills, which include lack in turn-taking and imitation skills; (c) conversation skills, which include trouble expressing themselves, deciding/knowing what to talk about and how to use appropriate body language; and (d) problem-solving skills, which include the inability to take decisions in social situations (Habash, 2014; Beaumont and Sofronoff, 2008; Kaat and Lecavalier, 2014).

2.3 Social Skills Training Interventions for Children with ASD

Social skills training (SST) interventions are a widely used approach for teaching and training social skills to children with ASD (O'Handley et al., 2015). SST interventions aim to teach children with ASD the necessary skills needed to communicate more effectively with others. Social skills training divides complex interactions into simple tasks (such as maintaining eye contact or responding with appropriate answers) towards achievable steps. Several literature reviews have been conducted to evaluate the effectiveness of these SST interventions in improving the social skills of children with ASD (Gresham et al., 2001; O'Handley et al., 2015; Barry, 2015). Peer mentoring, social skills group, video-modelling, social stories, picture exchange communication system, applied behaviour analysis, and occupational therapy are social skills training interventions for children with ASD.

2.3.1 Peer Mentoring

Peer mentoring or peer-mediated interventions (PMI) or peer buddies are promising intervention methodologies to improve the social skills of children with ASD (Chang and Locke, 2016; Płatos and Wojaczek, 2018). In this intervention, typically developing peers (e.g. classmates) are selected, trained and supervised to teach or support children with ASD (Bohlander et al., 2012). There are several approaches for applying PMI, such as peers who can model appropriate social behaviours or skills. Peers can act as intervention agents in school settings, and children with ASD can use peers to practise their newly acquired social skills or behaviours (Chang and Locke, 2016). Within the literature, PMI has demonstrated improvements in social, communication, academic, and play skills of children with ASD (Płatos and Wojaczek, 2018; Gunning et al., 2019; Bohlander et al., 2012; Chang and Locke, 2016).

2.3.2 Social Skills Group

Social skills group is a type of traditional intervention that allows children with ASD to practise social skills with each other or with typically developing peers regularly (Foden and Anderson, 2011; Lindgren and Doobay, 2011). The social skills lessons are taught by therapists or teachers either at the school or outpatient clinics and approximately comprise 4 or 5 children with ASD (Bohlander et al., 2012). These sessions teach skills to greet others, be friendly, join or initiate play with others, read non-verbal cues, and start and maintain conversations. Initially, the participants in social skills groups were only children with ASD, which limited the generalisation and the maintenance of the learnt skills (Bohlander et al., 2012). Involving typically developing (non-ASD) peers who are trained to encourage social engagement with peers with ASD was one of the factors that aid in the generalisation and maintenance of the skills learnt in social skills groups.

2.3.3 Video-Modelling

Video-modelling (VM) is a traditional social skills intervention used to teach children with ASD appropriate social behaviour. VM involves showing a video of a model demonstrating desired social behaviour to a child with ASD. VM either utilises peer or adult models or self-models (Delano, 2007). Video self-modelling demonstrates a video of the child himself/herself performing appropriate social skills. Adult/peer-modelling demonstrates a video of other children performing appropriate behaviour or skill. The child with ASD will be asked to imitate the skills or behaviours modelled in the video.

There are four styles of video-modelling. (a) Modelling with a video where the child watches a video of the target behaviour, then is asked to imitate the behaviour. (b) Feedback involves a video where the child watches himself/herself perform a scenario with appropriate and inappropriate behaviour. The situation is discussed with the therapist or trainer to make enhancements the next time. (c) A cue involves a video where an instructor gives the cue, and the child is asked to react based on that cue. (d) computer-aided video teaching which contains animation, sound, graphics, music, and texts in the same performance (Alzyoudi et al., 2015; Maione and Mirenda, 2006; Shukla-Mehta et al., 2010; Mechling, 2005).

Despite the empirical support for video-modelling, some drawbacks may reduce its utility and effectiveness. For example, VM studies mostly focused on teaching children with ASD daily living skills rather than social and safety skills, and studies need to be more focused on if the VM techniques are helpful for all age ranges and all types of autism. Video-modelling seems to be more effective when it is combined with other interventions, such as social stories.

2.3.4 Social Stories

Social stories are brief, personal stories describing specific social behaviour and skills, written for children to help them understand social situations (Foden and Anderson, 2011; Bohlander et al., 2012). The idea of using social stories to aid children with ASD was introduced by the educational consultant and former teacher Carol Gray in 1993 (Kokina and Kern, 2010). Gray and Garand (1993) set specific guidelines while creating a social story for children with ASD. The story needs to be in the first person, about a skill, behaviour, or activity and let the child know what, where, when, and why the situation or skill will happen, along with the expected behaviour and responses of the child with ASD. An example of a social story targeting greeting skill is "describing a child going up to another child, saying hello, and asking how the other child is doing and would explain why the child does this and the result" (Bohlander et al., 2012).

There are four basic approaches used in social story interventions for children with ASD (Gray and Garand, 1993; Sansosti et al., 2004). (a) For verbal children with ASD, the adult narrates the story two times to the child, then the child reads it back. Once familiar with the story, the child reads it once every day independently. (b) For children with ASD who cannot read, the adult produces an audio recording of the story, and the child is expected to listen to it daily. (c) A video recording of the story is produced. This approach combines social stories with video-modelling, where the story is videotaped. The story is read aloud on the videotape, with one page appearing on the screen at a time. (d) Presenting the story through a computer-based programme or assistive technology such as robots and virtual environments. Many studies were conducted to show the effectiveness of storytelling in enhancing the social communication skills of children with ASD (Ugwuanyi et al., 2018; O'Handley et al., 2015). The results from these studies are inconsistent as some studies indicate positive effects of using the storytelling intervention while other studies contradict its effectiveness (Lee et al., 2018; Scattone et al., 2006).

Storytelling intervention helps in reducing inappropriate behaviour rather than improving new skills (O'Handley et al., 2015). Using social stories as a separate intervention does not achieve the required effectiveness in enhancing the social skills of ASD because they read or listen to the story then try to mimic the situation without direct interaction with others (Lee et al., 2018; Daneshvar et al., 2019; Qi et al., 2018; O'Handley et al., 2015). The traditional social stories intervention does not show how to apply the skills in the real world and does not evaluate its effectiveness in real situations.

2.3.5 Picture Exchange Communication System

Picture exchange communication system (PECS) is an augmentative communication system developed by (Bondy and Frost, 1998). PECS is a teaching protocol that allows non-verbal children with ASD to communicate using the desired picture as sign language (Flippin et al., 2010). PECS aims to teach children with ASD functional communication and consists of six phases (Ganz and Simpson, 2004). (a) Phase I aims to teach the child how to communicate using single pictures for the items or activities he/she wants. (b) Phase II aims to generalise the learnt skill by using it in different situations with different people. (c) Picture discrimination is the aim of phase III, where the child learns to select from two or more pictures to ask for favourite things. (d) In phase IV, the child learns to construct a sentence using more than one picture. (e) In phase V, the child learns to answer questions using pictures such as what do you want?. (f) Phase VI aims to teach the child to respond to questions such as what do you hear? He/she learns to make a sentence starting with "I hear".

2.3.6 Applied Behaviour Analysis

"Applied behaviour analysis (ABA) is defined as the process of applying behavioural principles to change specific behaviours and simultaneously evaluating the effectiveness of the intervention" (Lindgren and Doobay, 2011). ABA is a widely used training intervention for children with ASD. ABA is beneficial in training children with ASD because it helps understand how behaviour works, how behaviour is affected by the environment and how learning takes place (Axelrod et al., 2012). That type of intervention should be provided under the supervision of a trained behavioural psychologist or behaviour analyst (Lindgren and Doobay, 2011). ABA therapy interventions can help increase language and communication skills, improve attention, focus, social skills, and memory, and decrease problem behaviours. It can also combine with other traditional training interventions.

2.3.7 Occupational Therapy

Occupational therapy (OT) aims to help children with ASD achieve or maintain their maximum independence level. OT interventions always occur in partnership with the child's family or teaching team and within a naturalistic environment (Makrygianni et al., 2018). Occupational therapy consists of two stages: evaluation stage and therapy stage (Autism Speaks, 2020). In the evaluation stage, the therapists assess the child's needs through observation of functional activities. In the therapy stage, the therapists develop a protocol or a training programme that meets the observed needs. Handwriting, fine motor, and daily living skills are the skills taught in these intervention sessions.

2.4 Conclusion

Autism spectrum disorder is an umbrella term that categories a group of disorders of brain development that include impairments in social and communication interaction, repetitive behaviours, and stereotyped patterns of interests and activities (Shindorf, 2016). According to estimation from CDC's Autism and Developmental Disabilities Monitoring (ADDM) Network, 1 in 54 children are identified with ASD (CDC, 2020). Children with ASD have difficulties in the theory of mind, emotions, and communication and social skills. Social and communication challenges can significantly affect their social life, including forming and maintaining relationships and functioning independently. Social skills training (SST) interventions teach children with ASD the skills necessary to navigate their social environment. Social skills training interventions such as peer mentoring, social skills group, social stories, video-modelling, picture exchange communication system, applied behaviour analysis, and occupational therapy were the early promising solutions for enhancing the social skills of children with ASD.

Despite the reported positive effect of SST , there are some barriers to accessing SST interventions for families with children of ASD. These traditional sessions are costly and require time-intensive training (Soares et al., 2021; Bekele et al., 2013). Another primary barrier to SST interventions is the shortage of trained therapists and facilitators (Soares et al., 2021; Dickstein-Fischer et al., 2018; LuxAI, 2021). Especially in countries like Egypt, parents emphasise the absence of early interventions and services for their children with ASD or their limited availability in private schools in the capital city (Gobrial, 2018).

The generalisation of the skills to the real-life context is the most salient criticism of the evidence for SST. One contributing reason is that most studies do not include methods for measuring the learnt skills' generalisation (Shindorf, 2016). Additionally, specialists emphasise the importance of real-life settings interventions for children with ASD in generalising the taught skills. However, there are barriers to implement these interventions in such settings, such as lack of time and lack of commitment (Shindorf, 2016). Furthermore, research indicates that there is an increase in the generalisation of skills when parents are involved in the intervention (Garbacz et al., 2016; Shindorf, 2016).

With the rapid growth of technology in recent years, assistive technology for cognition (ATC) can offer an alternative means to train the social skills of children with ASD through less expensive, repetitive, and more scalable options. It may also reduce the load and stress over the therapists and practitioners by reducing the burden of memorising the content and providing data regarding behavioural and communication skills. Technology may also mitigate the stress and discomfort children with ASD face during the traditional training sessions. With the technological advances, the sessions became more fun and engaging for children with ASD, resulting in increasing the learning outcome compared to traditional methods. The consistency that technological interventions provide creates predictability which gives children with ASD a sense of safety and makes them push themselves harder because they are not afraid of making mistakes. Finally, technological interventions benefit from the traditional SST interventions and combine the successful elements to develop a more engaging and effective environment for children with ASD.

3 Assistive Technology for Children with ASD - SoA

3.1 Chapter Overview

Chapter 2 provided a review of literature on the impairments of children with ASD in four key areas, including the theory of mind, emotions, and communication and social skills. It was also summarised the traditional social skill training interventions that have been used for training social skills of children with ASD.

In this chapter, assistive technologies for cognition that have been used in social skill training for children with ASD, virtual environments and social robots are reviewed. The virtual environments reviewed are classified based on their type: immersive and non-immersive environments. Social robots' studies have been divided into humanoid and animal robots. The chapter then follows with an overview of the assessment methods used by SoA to assess the social skills of children with ASD and evaluate the effectiveness of the technology interventions. The investigation into traditional social skill training (SST) interventions led to exploring the alternative interventions in training the social skills of children with ASD. This chapter aims to answer the following research question.

What are the alternative methods for improving the social skills of children with ASD?

3.2 Computer-Assisted Technology

Computer-assisted technology (CAT) and computer-based interventions (CBIs) are becoming promising tools for enhancing the social life of children with ASD. Children with ASD prefer using computers or watching television rather than communicating with other peers (Bekele et al., 2013; Carter et al., 2016). The inexpensive implementation, unlimited repetition, and safety are advantages of CAT and CBIs, which motivates their use in training children with ASD (Carter et al., 2016; Pour et al., 2018; Raj et al., 2018; Schuller et al., 2015).

Assistive technology for cognition refers to technologies that are used to enhance or facilitate cognitive function and include tools that aim to improve social participation and independent actions of individuals with cognitive disabilities (Desideri et al., 2020). Assistive technologies for cognition have been implemented using virtual environments (VEs) and robots to train and improve the social skills of children with ASD.

3.2.1 Virtual Environments Training Applications for Children with ASD

Virtual environments (VEs) are a computer-generated artificial environment that makes users feel they are part of these environments. The high level of interaction and immersion that VE offers

Figure 1: Immersive VE. (Rasmussen, 2016)

Figure 2: Non-immersive VE. (Poetker, 2019)

distinguishes it from other computer technologies. There are two types of virtual environments: Immersive VE as shown in Figure 1 and Non-Immersive (Desktop) VE as shown in Figure 2. In Immersive VE, the user is surrounded by the virtual environment via head-mounted displays (HMDs) or projection systems (CAVE). In contrast, Non-Immersive (Desktop) VE is a type of VE that uses computer monitors or TVs as displays, such as video games. Augmented Reality (AR), which can be considered as another type of immersive VE (Koushki et al.), combines interactive experiences in the real-world and 3D objects in a VE in real-time (Azuma et al., 2001).

In recent years, VE and virtual characters (VCs) have become the tools for training and supporting children with ASD (Abdelmohsen and Arafa, 2021). According to Nojavanasghari et al. (2017), VEs and VCs can enhance the social skills of children with ASD by simulating real scenarios in daily life and will help them practise social skills. Children with ASD are attracted to systems that contain animations and sounds, which help to reduce their stress and engender a sense of safety (Zhang et al., 2018). Using VE as a training tool for children with ASD has several advantages, such as practising in a safe environment, gradually increasing the level of complexity, decreasing stress, simulating realistic scenarios, providing immediate feedback, and facilitating the repetition of tasks (Bozgeyikli et al., 2018; Bozgeyikli, 2016; Bekele et al., 2013; Nojavanasghari et al., 2017).

Virtual characters have been utilised in virtual social scenarios to support children with ASD. For instance, some social skills training scenarios involve finding a place to sit in a crowded cafe or a bus, initiating interactions in a classroom or a playground (Lorenzo et al., 2013; Cheng et al., 2015) and shopping situations (Stewart Rosenfield et al., 2019).

3.2.1.1 Immersive VE Training Applications

As discussed earlier, immersive VE is a type of VE where the user is entirely surrounded by images, sounds, and 3D objects either via HMDs or CAVE. The number of studies that used immersive VE in training children with ASD is low compared to those using non-immersive VE. This is due to the high cost of immersive virtual reality equipment.

Jarrold et al. (2013) developed an immersive virtual environment to study the social attention of children with HFA. In this VE, HMD was used to display a virtual classroom where the child interacts with nine avatar peers (virtual students). The virtual avatars were programmed to blink and display head motions to enhance the level of immersion. 37 children with HFA (aged 8-16 years) and 54 typically developing children (Non-ASD) participated in the study. The study found that children with ASD looked less frequently to the avatar students in the classroom while talking than the typically developing children.

Lorenzo et al. (2013) designed an immersive virtual environment for teaching social skills to children with HFA. The designed virtual environments included a classroom, school playground, and bedroom. In the virtual playground and virtual classroom, the child interacts with virtual teachers and virtual students. The facial expressions, voice, and gaze of the user were monitored for evaluation. 20 children with HFA (aged 8-15 years) were recruited for this study for ten months (16 boys and 4 girls). The participants' tutors and teachers also participated in the study to observe and record the changes in their skills and behaviours after receiving the training sessions. The results show that participants were engaged and motivated by the virtual avatars. Additionally, some skills generalised to school context, as they applied the instructions practised in the virtual classroom into the real-life classroom.

Cheng et al. (2015) conducted a preliminary study on the use of an immersive virtual environment (3D-SU) to train children with ASD in different social situations. The system consists of two social situations: a virtual classroom and a bus stop (e.g., chatting with the bus driver, raising hands before asking questions, hiding someone's things). Both environments were presented via HMD. Several virtual characters were designed to interact with the child throughout the scenario, such as the bus driver and students. Three boys (aged 10-13 years) were involved in the study. Interaction with the system involves three stages: baseline, intervention, and maintenance. The study lasted six weeks and involved three sessions for each phase, which took approximately 30 to 40 minutes to measure the participants' behaviour changes. The sessions were video-recorded and audio recorded for analysing the child's responses and recording the target behaviours. The results show an improvement in the participants' ability in initiating social interaction during the interventions. The results suggest that fully immersive VE was effective in improving such skills. A limitation of this study is that it did not assess the generalisation of the taught skills.

Another study by Lorenzo et al. (2016) was developed to train the emotional skills of children with ASD. 40 children with ASD (aged 7-12 years) participated in the study for four months and included 29 boys and 11 girls. 20 participants were assigned to the intervention group and the rest to the control group. Ten social situations were designed to train the emotional ability of the participants. The virtual environments included a classroom, park, playground, and birthday party. Data was collected through teacher interviews and by observing facial expressions (section 3.4) to evaluate the child's engagement during the sessions. The results show an improvement in the participants' emotional competencies, and they interacted with the designed avatars in the virtual environments. The teacher's interviews show that participants' emotional ability in the real classroom improved, indicating that skills were transferred from the virtual environment to real life.

Another immersive virtual environment was developed by Stewart Rosenfield et al. (2019) to practise the conversation skills of children with ASD. The virtual fish pet shop was implemented in Unity 3D. A virtual human avatar (Bob) was deployed to interact with the children and help them practise joint attention and some non-verbal skills such as responding to waving. The virtual environment is similar to cartoon animations. Data was collected through participants' interviews and observations (section 3.4) to evaluate the system effectiveness. Two children with ASD (aged 6 and 7 years) participated in this study, a boy and a girl. The results show positive feedback from the participants, and they learnt how to interact with the environment quickly. As the sample size was small, and the fact that participants received only one training session, it was not sufficient to draw any conclusions regarding improving the participants' conversational skills.

Herrero and Lorenzo (2019) developed an immersive virtual school and playground via HMD for improving and training the emotional and social skills of children with ASD. Unity 3D game engine was used. In the VE, the children interact with several 3D characters, particularly with a teacher and 6 children, 3 boys and 3 girls. The Wizard-of-OZ was used where the researchers controlled the avatars' answers from a limited set of pre-defined answers created. The gaze of the users was monitored and used for the evaluation. 14 children with ASD (aged 4-15 years) participated in this study, including 13 boys and 1 girl. Seven children were assigned to the intervention group and the rest as a control group. The researchers observed that the participants show significant improvements in social and emotional skills, as well as non-verbal communication. The generalisation of the learnt skills was not assessed.

An augmented reality self-facial modelling system was designed by Chen et al. (2015) to enhance the emotion recognition and expression of three adolescents with ASD (aged 10-13 years). The authors took each participant's front and side facial images to create six facial expressions based on the Facial Action Coding System (FACS) models in 3D Max. Unity 3D game engine was used to animate these generated models. Vuforia and Qualcomm were integrated into Unity to build the AR platform. The system uses an augmented mirror where the participants could see themselves with 3D virtual facial expressions. The session starts with the therapist introducing a scenario about social communication. The participant is expected to select and wear a mask with the appropriate facial expression for this scenario. A real-time 3D facial expression appears in the augmented mirror over the participant's body. The experiment consists of three phases: baseline, intervention, and follow-up to measure the enhancement of the participants after using the intervention. The results show that the participants' performance in the baseline was low. It was improved in the intervention phase and slightly lowered in the follow-up phase. The study did not assess the generalisation of the taught skills.

Lee et al. (2018) designed an AR system based on a social story to train the greeting skills of children with ASD. Vuforia toolbox was used to create the environment and to make it more realistic. The system has 20 greeting scenarios from daily life situations at school, home, and in the community. Three children with ASD (aged 7-9 years) participated in this study for six weeks, with two sessions per week. The study went through three phases: the baseline phase, the intervention phase, and the maintenance phase. The results show that all children had lower scores in the baseline. The performance significantly increased in the intervention phase and retained it during the maintenance phase. Based on the results, the percentage of the error rate between the baseline phase and the maintenance phase decreased by approximately 50%.

Recent research was conducted by Lorenzo et al. (2019) to evaluate the use of AR training programmes for improving the social skills of children with ASD. 10 boys and 1 girl with ASD (aged <6) participated in this study for 20 weeks, with two sessions per week. AR mobile application (Quiver Vision), which has different social situations, was used. The authors split the participants into two groups: the control group (5 children) and the experimental group (6 children). The results show no significant changes before and after the AR intervention, although some items as flexibility and imitation showed a slight increase.

3.2.1.2 Non-Immersive VE Training Applications

As mentioned earlier, desktop or non-immersive VE is a type of VE that provides the user with a computer-generated environment. Users do not experience immersive involvement. They use input devices such as the keyboard, mouse and controllers to interact with this environment. Non-immersive VE is used widely to help children with ASD with their social, safety, and life skills.

Abirached et al. (2011) designed virtual 3D characters (girl avatar, boy avatar, and an alien) in a serious game (LIFEisGAME) to teach children with ASD to recognise facial expressions. In the beginning, the child is asked to choose an avatar. The game consists of four modes. In the first mode, the child is asked to recognise the emotion expressed by the 3D character. The second aims to build a face where the child is asked to drag-and-drop facial features to build an emotion. In the third mode, the child's facial expressions are replicated by the 3D character, and then the child is asked to mimic the 3D character's facial expression. In the last mode, the child is asked to identify the emotion from a social situation in a story told by the 3D character. The researchers ran a preliminary study to evaluate the game design. Nine children with ASD (aged 4-11 years) participated in this evaluation phase, including 7 boys and 2 girls, with their parents. The data was collected by interviewing the children and their parents and by recording the trial session (section 3.4). The results suggest that children with ASD prefer different avatars, such as humans, animals or aliens, although the avatars need to be in a cartoon shape. The study did not report any improvements in the participants' emotion recognition and expression skills.

Alcorn et al. (2011) developed a non-immersive virtual environment (ECHOES) to allow children with ASD to understand and explore interaction skills. The VE utilised a child-liked virtual character (Paul) to interact with the children. Children are expected to help Paul to collect objects by following the character's gaze and/or pointing. 32 children with ASD (aged 5-14 years) participated in the study, including 29 boys and 3 girls. Each child attended one experimental session. The session was recorded for further analysis. As a result of this trial experiment, the researchers affirmed that the virtual character successfully engaged the participants. However, the study did not report any changes in the participants' behaviours or skills because of the short training session.

Bekele et al. (2013) developed a VE-based facial emotional recognition system that monitors eye gaze and physiological signals of adolescents with ASD. The system consists of eye tracking, physiological monitoring, and the VE task presentation engine. The Tobii software development kit was used in the eye-tracking application to monitor the participants' eye gaze. Biopac (SDK) and BioNomadix wireless physiological acquisition modules were used in the physiological monitoring application. In the VE task presentation, Mixamo and Autodesk Maya were used to generate the characters and import them into the Unity 3D game engine for the final presentation. The animated human-like characters presented seven emotional expressions with four intensities for each emotion in different scenarios at school. 10 adolescents with HFA (aged 13-17 years) and 10 typically developing adolescents (TD) as a control group participated in this study. This experiment shows that participants with ASD pay attention to irrelevant areas of the face, such as the forehead, rather than the eye and mouth areas. Adolescents with ASD looked 11.32% more to the forehead and 12.1% less to the mouth region than the control group, and 6.41% more to the non-facial areas and 8.25% less to the face area. They also struggled to differentiate between the extreme and low intensities of emotions.

Serret et al. (2014) designed a serious game (JeStiMulE) that teaches children and adolescents with HFA emotion recognition skills in a 3D environment. The game targets training emotion recognition skills through facial expressions, body gestures, and social situations. Six basic emotions (happy, sad, fear, surprise, anger, and disgust), one feeling (pain), and two complementary expressions (neutral and funny) were expressed by human-like characters in this game. 33 children and adolescents with HFA (aged 6-17 years) participated in this study, including 31 boys and 2 girls. In this study, colours are used to represent the basic emotions: happiness= yellow, anger= red, sad= light blue, fear= green, surprise= dark blue, and disgust= purple. In the beginning, the participants choose and personalise their avatars. The VE includes a garden, restaurant, shop, theatre, and square where the participants control their avatars in different social situations. Once the situation was finished, the participant had to identify the emotion expressed in that situation. The results show that 73% of the participants successfully completed all the game levels (recognising emotions from avatar faces, gestures, and in social contexts). After the training sessions with JeStiMilE, participants were more accurate in emotion recognition tasks.

After the training sessions, the participants were asked to identify the emotions from photographs. The results indicate a potential for generalising the skills learnt in the VE to real life.

A virtual social cognition training environment (VE-SCT) was designed by Didehbani et al. (2016) to train the social skills of children with HFA. The platform contains various social contexts and varying complexity suitable for different ages. 30 children with HFA (aged 7-16 years) enrolled in the study for five weeks with two sessions per week (1h per session). Second Life version 2.1 platform, 3D virtual environment public software, was used in the intervention, including several scenarios, such as classroom, playground, central park, restaurants, and stores. The participants are represented in the virtual environment as avatars. Two professional clinicians were involved in the study (Coach, Confederate). The coach describes and moderates the sessions, while the confederate is portrayed as another child in the VE and interacts with the participants. Participants navigate in the VE by using a mouse or a keyboard. Results show improvement in the participants' emotion recognition skills but not in their social attribution skills. The generalisation of the taught skills was not assessed.

Nojavanasghari et al. (2017) designed a non-immersive virtual environment to enhance the social skills of children with ASD. The study targets three social skills; initiating interaction, understanding emotions, and learning turn-taking. The designed avatars had a child-like appearance. The VE consists of two stations; the child's station and the interactor's station. In the child's station, the child interacts with the avatars. The child's eye contact and facial expressions were monitored to be used as feedback about the child's performance during the interaction with the environment. In the interactor's station, the interactor controls the avatars according to the child's affective state that was gathered in the child's station. The researchers did not conduct a case study and did not report any information regarding the effectiveness of the VE.

A serious game (Pico's Adventure) in the form of a virtual environment was designed by Malinverni et al. (2017) to promote social interaction in children with ASD. The game utilised an alien avatar (Pico) to interact with the child. Ten children with HFA (aged 4-6 years) participated in the study for one month, with a session per week. Data was collected through live and video observation (section 3.4) to record and rate the child's behaviour during the sessions. The participants are expected to interact with Pico by offering food and fixing its spaceship. The child is expected to play with an adult in some game levels, such as the parent or the researcher. The findings of this study show the acceptance of the game design to the participants. Additionally, it was observed that the participants' repetitive behaviours were reduced by comparing the first and the last session that indicates the child's engagement with the VE. The study did not assess the child's behaviours in the baseline sessions; thus, the researchers could not report any changes regarding the target skills. Furthermore, they could not evaluate the effectiveness of the proposed game in generalising the target skills to the real world.

In a recent study, Moon and Ke (2019) designed a desktop-virtual environment to train the social skills of children with HFA. 15 children with HFA (aged 10-14 years) were recruited in this study, including 13 boys and two girls for 8-14 weeks. The VE was developed using Open-Simulator software. Social scenarios that simulate social problems were designed to be similar to real-life problems, which may occur in the real world. The target skills in this study were negotiating, initiating interaction, responding to interaction, and collaborating. The participant is expected to engage in social tasks in different social scenarios, such as (a) interviewing a person to be hired; (b) working as a waiter in a restaurant; (c) designing a building by client's request; (d) resolving a bullying situation in a park. Three participants completed all social scenarios, while two participants achieved scores lower than the expected criterion. The majority of participants found it hard to initiate interactions, specifically as an interviewer. Furthermore, the study emphasises the importance of using visual stimuli in the VE as some participants struggled to understand the written instructions.

A study was conducted by Ghanouni et al. (2019) to validate the effect of social stories in training emotion recognition and perspective-taking skills of children with ASD. The target participants in this study were the parents and clinicians working with children with ASD. This study aimed to evaluate the content of the virtual environment to train the target skills of the children. The participants were asked to rate the social scenarios designed for VE training. Each scenario describes a short emotional story told by an avatar. The goal of each scenario was to take the avatar's perspective and respond with the relevant emotion. The results indicate that presenting social stories in animations motivates children with ASD and makes the learning process enjoyable. Furthermore, using pictures while narrating a social story could facilitate communication among children with ASD and increase their engagement with the virtual environment. Parents and clinicians highlight the importance of the level of difficulty in the stories and the intensity of emotions. They also suggest teaching children with ASD calming strategies and appropriate responses to emotions. Additionally, participants indicate the importance of teaching children the reasons behind the relevant emotion and behaviour.

Charlton et al. (2020) conducted a study to evaluate the effect of using a digital avatar in delivering social lessons to children with ASD. Five children with ASD (aged 8-10 years) were recruited for this study. Data was collected through video recording, parent questionnaire, and live observation (section 3.4). The 3D character chosen for this study was a colourful fish (Marla) with facial expressions, similar to a character from the cartoon movie Finding Nemo. The target skill for this study was initiating conversations. The results show that the participants' initiating conversation skills improved by comparing the baseline session to the intervention session. Additionally, the participants generalised the skills learnt in the intervention session to their daily life. They started to talk more with their peers. An additional finding in this study is that the participants enjoyed interacting with Marla. This result indicates that using popular cartoon characters increases the engagement of children with ASD with the interventions.

A recent study was carried out by Ke et al. (2020) to explore the usage of a desktop virtual learning environment in enhancing the social skills of children with ASD. Seven children with ASD (aged 10-14 years) took part in this study, six boys and one girl. Different social scenes were designed, such as a school, museum park, neighbourhood, and public facility using OpenSimulator for practising everyday social skills. The participants control human avatars in the scenes to interact with puppeteered social characters within the social context. Data was collected through observation (section 3.4) and pre and post-assessments (section 3.3). The findings show improvement in the target social skills from baseline to the intervention phase. However, the study did not measure the generalisation of the taught skills.

3.2.1.3 Critique of the Research Outcomes

The development and the use of virtual environments for ASD interventions are on the rise, and their potential benefit for enhancing the social skills of children with ASD is numerous. After reviewing the studies that aimed at evaluating the use of VEs in enhancing the social skills of children with ASD, there appear to be clear limitations. A significant finding is the use of small sample sizes. To highlight this, only 5 studies of the 18 studies reviewed in this chapter had a sample size greater than 20 as shown in Table 1. This finding is consistent with previous systematic reviews. For example, in Miller and Bugnariu (2016) meta-analysis, only 7 of the 24 studies included had a sample size above 20. Similarly, Mesa-Gresa et al. (2018) reported that only 6 of 31 studies reviewed had participants over 20. Furthermore, in Khowaja et al. (2020) systematic review about the use of AR environments for children with ASD, 57% of the reviewed studies had less than 10 participants, whereas the remaining 43% of the studies had more than 10 participants. Only two of these studies had more than 21 participants. These statistics reflect the difficulty in recruiting individuals with ASD, especially if they are children. It is well-known the difficulty of recruiting large samples when the study's aim is related to psychology (Mesa-Gresa et al., 2018).

The generalisation of the acquired skills in such environments remains an open question. A consistent limitation noted in many of the reviewed studies and literature reviews is that these studies either did not measure the generalisation of the acquired skills or did not provide the generalisation results (Vasquez et al., 2015; Didehbani et al., 2016). To highlight this, only 4 studies of the 18 studies included in this chapter have assessed the generalisation as shown in Table 1. This finding is in line with Khowaja et al. (2020) systematic review, where only 1 study of 30 studies included has conducted a generalisation test. However, the authors of this study did not provide the results of the generalisation test.

Furthermore, the type of technology used in virtual and augmented environments affects the availability of these interventions. The cost of these technologies is a potential barrier to the widespread use of immersive VE interventions. For instance, projection-based technology is costly and challenging to set up, as it is set up in a dedicated classroom. Therefore, the number of desktop virtual environment studies is more than immersive VE and AR interventions. Thus, desktop VEs are preferred over immersive and AR environments. Additionally, many children with ASD face difficulty wearing or interacting with the specialised virtual headsets. According to Malihi et al. (2020); Bradley and Newbutt (2018), the use of HMDs has been associated with negative effects, including cyber-sickness, headache, and dizziness. Nonetheless, the reviewed studies using immersive VE and desktop VE reported positive learning outcomes to suggest that the level of immersion may not be critically important in the effectiveness of these interventions. However, in the absence of research comparing immersive and non-immersive (desktop) VEs, it is not possible to draw a clear comparison between these VEs (Berenguer et al., 2020).

Finally, there are concerns that VEs could cause more social isolation of children with ASD and increase their avoidance of the interaction with the real-world (Berenguer et al., 2020; Malihi et al., 2020; Bradley and Newbutt, 2018). This highlights the importance of using these interventions in the presence of parents.

3.2.2 Robot-Assisted Training for Children with ASD

Children with ASD react better to robots than humans because they are predictable and straightforward (David et al., 2020; Zheng et al., 2017). It was noticed that robots could capture the attention of children with ASD (Pour et al., 2018; Taheri et al., 2015; Marinoiu et al., 2018). The social robot has been widely used in ASD therapy. Social robots are controllable and predictable, as well as they can be programmed to repeat the same task several times, so they provide clear and structured information. Additionally, they can reduce the anxiety of children with ASD (So et al., 2018b). According to Zheng et al. (2017), where a person interacts with a child with ASD in two different ways, the first way was the person dressed and acted like a robot, while in the second way, the person acts normally. This study shows that the child interacted and spoke more with the person when dressed as a robot.

There has been increased interest in effective robots for educating and training children with ASD in recent years. According to Thill et al. (2012), the aim of Robot-Assisted Therapy (RAT) is to improve the communication and social interaction skills of children with ASD and help them in demonstrating and perceiving emotions. Some therapists use the robot in ASD therapy to enhance children's emotion recognition skills (Miskam et al., 2014; Marino et al., 2020; Salvador et al., 2015). For instance, the therapist shows the child pictures of different emotions, then the robot expresses these emotions, and the therapist asks the child to recognise the robot's emotion (Rudovic et al., 2018). Another example is enhancing the theory of mind in children with ASD through storytelling techniques, where the therapist starts a new task by telling a story about the robot and asks the child about the robot's feeling at the end of the story (Rudovic et al., 2018; Costa et al., 2014).

Social robots can be classified into three main categories: i) Anthropomorphic: Robots with some human features such as NAO (Softbankrobotics, 2017), KASPAR (UniversityofHertfordshire, 2005), ZENO (HansonRobotics, 2005) and QTRobot (LuxAI, 2017);

ii) Non-anthropomorphic: Robots with biological features but do not have human characteristics such as robotic animals such as MiRo (Consequentialrobotics, 2020), AIBO (Sony, 2006), and Kasha (Stanton et al., 2008); and

iii) Non-biometric: robots that do not resemble any biological creatures such as Pekee and Roball (Valadão et al., 2016; Zheng et al., 2017).

3.2.2.1 Anthropomorphic Robot "Humanoid Robot"

Humanoid robots have a human-like appearance. They have a humanoid head, body, and arms. Humanoid robots can move different parts of their body, and some of them can demonstrate facial expressions.

3.2.2.1.1 NAO Robot

NAO is a humanoid robot created by SoftBank Robotics in 2008, as shown in Figure 3. It has been used as an assistant by companies and healthcare centres. It can produce a wide range of body gestures. The NAO robot contains 25 degrees of freedom from 15 joints and actuators. NAO has been widely used in ASD training.

Figure 3: NAO Robot.
(Softbankrobotics, 2017)

Greczek et al. (2014) conducted a study to train the imitation skills of children with HFA. 12 participants with HFA (aged 7-10 years) participated in the study to play a "Copy-cat" game with NAO robot for 2.5 weeks, with 2 sessions per week. During the game, NAO asks the child to imitate 25 different hand movements. For example, NAO raises both arms and asks the child "Can you copy me?". If the child's action is correct, NAO flashes its eyes green, nods, and says "very good" otherwise; NAO gives prompts such as "Are you sure that's right?", then says "let's try another one" and moves to the next arm movement. The results show that participants' imitation skills improved throughout the sessions. Additionally, it was observed that providing feedback kept the children motivated in the sessions.

Miskam et al. (2014) used NAO in a study to teach children with ASD how to express emotions using body language. They also developed an android application that everyone can easily use to control the NAO robot. NAO shows emotions using body postures and colours, as it does not have facial expressions. NAO expresses nine emotions: fear, happiness, sadness, hungry, angry, shyness, surprise, loving, and tiredness. The authors experimented with 1 typically developing child (aged 6 years) and a therapist. The participant engaged well with NAO and was able to identify the correct emotions. The therapist suggests that NAO has a good appearance and could facilitate the learning process. Additionally, the therapist found the application very easy to use.

A study was conducted by So et al. (2018b) to teach children with ASD to recognise and produce gestures using NAO. 45 participants (aged 4-6 years) were recruited, including 30 children with ASD and 15 typically developing children. The participants with ASD were divided into two groups: intervention and control groups. NAO was programmed to produce 14 intransitive gestures commonly used in daily life, such as hello, not allowed, wrong, and awesome. The results show that children with ASD who received the intervention training produced intransitive gestures more accurately than those in the control group. They also found that robot-assisted intervention can develop intransitive gestures in young children with ASD and prevent its delay. Furthermore, the participants generalised the acquired skills in a novel setting with a human.

Feng et al. (2018) used NAO to present a control architecture that increases the robot autonomy in ASD training. The intention behind increasing robot autonomy is to reduce the workload of humans as therapists, parents, and educators. Two children with ASD (aged 6 and 4 years) participated in the study, including a boy and a girl. The target skill was imitation skill. NAO was programmed to ask the child to imitate its action. The participants engaged with NAO and imitated its hand movements. However, the generalisation of the skills was not assessed.

Marinoiu et al. (2018) presented a study to investigate the effectiveness and feasibility of using NAO. 14 children with ASD (aged 4-8 years) enrolled in the study, seven children in an experimental group and seven in a control group. NAO was used as a co-therapist in the sessions. The sessions lasted for 12 weeks, with two sessions per week. The study aimed to train the emotion recognition skill of children with ASD through a set of emotion recognition games presented by NAO. The sessions were video recorded for further analysis. In the pre-intervention phase, none of the groups was able to recognise the shame emotion. In contrast, the experimental group recognised the five basic emotions more efficiently by 48% as well as the shame emotion in the post-test (section 3.3). A limitation of this study is that it did not assess the generalisation of the skills.

So et al. (2019) conducted a study to evaluate the learning outcomes in children with LFA from robot-based intervention compared to human-based intervention. 23 children with LFA (aged 6-12 years) participated in the study, including 20 boys and 3 girls. Two NAO robots were programmed to show 14 gestures such as hungry, yes, and hello. One robot acts (Mary), and the other acts (Sally). In the robot-based training sessions, Sally and Mary engage in a conversation while demonstrating different gestures. Each time the child is asked to imitate the gestures. The same procedure is applied in the human-based training sessions; however, the child watches two humans engage in a conversation. The results indicate that the participants had more eye contact with the robot compared to a human. No significant differences were found in the imitation skills of the participants between the robot-based intervention and human-based intervention.

Cao et al. (2019) presented the DREAM project, which aims to develop a new generation of robot-assisted therapy, RET (Robot Enhanced Therapy), to meet the clinical and technological perspectives. Roboticists, psychotherapists, cognitive scientists, computer scientists, and ethicists participated in the project development to allow the robots to behave autonomously while ensuring safe and ethical behaviours. Imitation, turn-taking, and joint attention were the target skills for the DREAM project based on therapeutic interventions (applied behaviour analysis). Facial expressions, gaze, vocal prosody, and movements were monitored to understand the child's robot interaction. The project seeks to determine how much RET can enhance children's skills with ASD compared to standard interventions. The experimentation was divided into two phases: one using NAO robot under a WoZ control system and using it within a supervised autonomous system. WoZ control system means that the user remotely controls the robot. Both phases were compared at the end with standard human treatment (SHT) conditions. 38 children with ASD (aged 3-5 years) participated in this study (11 for the first phase and 27 for the second). The results of the first phase depended on the being learnt skills. During turn-taking, WoZ achieved better results than SHT. WoZ and SHT showed similar results for the joint attention task as well as for the imitation task. The second phase is currently being done to get better answers about the advantage of supervised autonomous systems in ASD therapy.

Peca et al. (2020) conducted a study to investigate the child's awareness when being imitated by NAO compared to being imitated by a human. NAO and a human mirror the child's movement: rising left/right arm vertically, rising left/right arm horizontally, rising both arms vertically, and rising both arms horizontally. Four children with ASD participated in the study for four weeks. During the experiment sessions, the child interacts directly with the robot, and the operator observes the child's movements and controls the robot's movements using a Kinect sensor. After analysing the gathered data, some children were smiling when they understood the robot imitates their moves. This finding indicates that NAO has a positive effect on the participants and succeed in capturing their attention. Additionally, the participants interacted more with NAO compared to the human. The generalisation of the skills was not measured. Furthermore, the study did not report the characteristics of the participants, such as age or gender.

Another study was conducted by So et al. (2020) utilises NAO to train the joint attention and play skills of children with HFA through a play drama intervention. Two NAO robots were programmed to act in dramas as main actors. 23 children with HFA (aged 4-6 years) participated in the study, including 20 boys and 3 girls. The participants were divided into intervention group (N=12) and control group (N=11) for 9 weeks. Three drama scripts were designed for this study: Tourist and Tour Guide, Farmer and Butterfly, and Doctor and Patient. In the training sessions, NAO robots act as the main characters in the drama while the child is watching them. The child is then asked to act as one character and participate in a role-play with the robot. After finishing the drama, they swap the roles and participate again in a role-play. At the end, the child is asked to participate in a role-play with a human to evaluate whether the child could generalise the skills learnt in the training. The results show that participants who received the training session improved their joint attention and play skills compared to those in the control group. Additionally, the participants engaged and showed improvement when participated with a human, which indicates that they generalised the skills learnt in the training sessions.

3.2.2.1.2 KASPAR Robot

KASPAR is a child-sized humanoid robot that has been used in various human-robot interaction studies. The Adaptive Systems group has developed it at the University of Hertfordshire as shown in Figure 4. KASPAR has three operation modes: automatic behaviour, controlled operating mode, and semi-autonomous mode (UniversityofHertfordshire, 2005).

Figure 4: KASPAR.
(UniversityofHertfordshire, 2005)

Costa et al. (2015) used the KASPAR robot to interact with children with ASD through a robot-assisted play scenario. The scenario teaches children the human body parts and facilitates interaction between the child and the researcher. Eight boys with ASD (aged 6-10 years) participated in seven sessions. The study was divided into four phases: familiarisation, pre-test, training, and post-test (section 3.3). Data was collected through interviews, questionnaires, and video observation (section 3.4) to evaluate the interaction between the child, robot, and the experimenter. The results show that KASPAR facilitated the interaction between the children and the experimenter as a social mediator. Furthermore, the participant's ability to identify different human parts increased after completing the seven training sessions. The most promising finding is that the participants could generalise some gained skills during the training sessions to their classroom.

Huijnen et al. (2017) conducted a study to understand the essential requirements while implementing robot interventions for children with ASD and describing the intervention template that allows professionals to create new interventions. Professionals, therapists, ASD teachers, assistants at special needs schools, and ASD parents participated in the study. KASPAR robot was used in this study. The study discusses the requirements related to the robot appearance, behaviours, and attributes. Some professionals indicate that the KASPAR's voice could be computerised voice and friendly. The talking speed must be slow, and the sentences must be short and understandable. They also noticed that when KASPAR moves, they could hear the sound of the motors, which might distract children with ASD. Professionals also indicate that they need the training to be able to operate and interact with KASPAR. They also developed some social scenario interaction and templates of intervention details to be used in new interventions. Furthermore, professionals and therapists suggest that the intervention sessions should be short to match the attention span of children with ASD (max 30 mins).

Zorcec et al. (2018) presented a case study using KASPAR to evaluate the interaction between children with ASD and the robot. Two children with ASD participated in the study. Learning emotions, greeting skills, imitation skills, and learning animal sounds were the target skills of this study. It was observed that the participants showed aggressive behaviour towards KASPAR (such as poking its eyes) for some sessions, then that behaviour gradually declined throughout the sessions. After ten sessions of the greeting skill task, the children started using the greeting skills in their daily lives, indicating that KASPAR was effectively successful in generalising the taught skills. They also got engaged in learning emotions and imitation tasks. Parents reported that their children learnt to differentiate between happy and sad.

A recent study was conducted by Wood et al. (2019) to improve the visual perspective-taking (VPT) of children with ASD. VPT is the ability to see the world from other's points of view. 12 children with ASD (aged 6-14 years) were recruited for this study, including 7 boys and 5 girls. The researchers used a screen to display what KASPAR sees. The environment layout consists of a camera, screen, KASPAR, a cube with different animal pictures on its faces, researcher, and the child. The screen was placed next to KASPAR, and the child sits in front of them in order for him to see what KASPAR sees. In the training sessions, the participants played 9 games with KASPAR. For example, the child picks a random animal toy, and KASPAR makes the sound of that animal. The games emphasise that the face of the cube towards the child differs from the face of the cube facing KASPAR. An immediate post-test (section 3.3) without KASPAR was conducted to evaluate whether the participants could generalise the VPT skills. The results of the post-test indicate that 7 of the participants generalised the taught skills. Furthermore, it was observed that the participants engaged in most of the games with KASPAR.

Another study by Huijnen et al. (2021) was utilised KASPAR to evaluate its impact on children with ASD. Nine children with ASD (aged 8-12 years) participated in four sessions, including 8 boys and 1 girl. Each child receives 2 sessions with KASPAR and 2 sessions with a teacher. All the sessions were video-recorded to evaluate the child's interaction with KASPAR and the teacher. The sessions' goal was making contact, which means the child is expected to make a conversation and asking questions. The results indicate that the participants concentrated more on KASPAR compared to the teacher. They engaged more in the sessions with KASPAR than with the teacher. These findings emphasise that KASPAR has a positive effect on children with ASD.

3.2.2.1.3 Zeno Robot

Zeno is a humanoid robot developed by (HansonRobotics, 2005); called Milo as shown in Figure 5. Zeno can interact with people using facial expressions and vocals. It also has several gestures which could be used to attract the child's attention.

Figure 5: Zeno.
(HansonRobotics, 2005)

Costa et al. (2014) presented a study to develop emotion recognition skills of children with ASD through imitation and storytelling scenarios. Zeno was programmed to express different facial expressions. Three children with HFA (aged 12-15 years) participated in three sessions, including 2 girls and 1 boy. In the imitation scenario, Zeno expresses emotion, and the child is expected to imitate the same facial expression. Zeno narrates a story in the storytelling scenario then asks the child to show a racket with an image that identifies Zeno's feelings. Data was collected through video observation (section 3.4) to record the child's behaviour and skills. The experiment results show that the children understood the social stories and answered with the relevant emotion. In the imitation scenario, the participants imitated the robot's facial expression with 100% accuracy; however, the generalisation of the acquired skills was not assessed.

Salvador et al. (2015) conducted a study that utilises Zeno to examine the ability of emotion recognition in children with ASD compared to typically developing children. Zeno was programmed to express six basic emotions using facial expressions and body gestures. 11 children with HFA (aged 7-11 years), including 9 boys and 2 girls, and 11 typically developing children (TD), including 6 boys and 5 girls, were recruited for this study. In the emotion recognition game, Zeno expresses an emotion, then asks the child to indicate the name of this emotion. The results show that children with ASD had no significant impairments in recognising facial expressions compared to their typically developing peers. Additionally, a significant improvement was found for recognising DISGUST and FEAR emotions in children with ASD compared to TD after adding gestures. This finding indicates that incorporating gestures with facial expressions facilitates the learning process of emotions.

A recent study was conducted by Soares et al. (2019) to evaluate the influence of Zeno in improving the emotion recognition and expression skills of children with HFA. 45 children with HFA (aged 5-10 years) were divided into 3 equal groups, including 36 boys and 9 girls. The first group performed game scenarios with Zeno. The second group performed the game scenarios without Zeno (with a human partner). The last group did not receive any intervention. The game scenarios were divided into 3 tasks: recognise, imitate, and storytelling. Significant improvement in recognising and expressing facial expressions of Zeno's group was found compared to the other groups. Additionally, 30% increase in the participants' emotion recognition skills from social stories. The participants in Zeno's group generalised the learnt skills.

3.2.2.1.4 QT Robot

QTrobot is a child-sized humanoid robot as shown in Figure 6. It was built and designed by LuxAI to assist children with ASD in learning new skills. QTrobot has a screen that displays facial animations allows animation faces so it can demonstrate emotions. It can also perform different body gestures with its upper body.

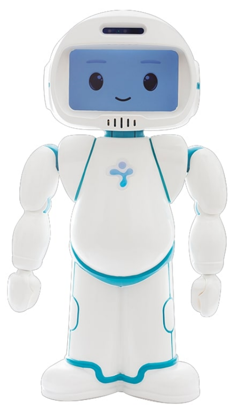


Figure 6: QTrobot.
(LuxAI, 2017)

Costa et al. (2017) conducted a study using QTrobot to teach children with ASD emotional ability skills. 40 children with ASD (aged 5-12 years) were recruited in this study. The sessions between the children and the robot were controlled by a therapist using QT's tablet application. The sessions start with a brief introduction about the sessions' objectives and the training programme and summarise what was learnt at the end. Data was collected through observations and questionnaires (section 3.4) to assess the intervention.

Another study was conducted by Costa et al. (2018) to investigate the effectiveness of QTrobot through interventions with children with ASD. 15 children with ASD (aged 4-14 years) participated in this study. The study duration was only two hours. The study aims to compare the interaction between the child with a human and with a robot. The intervention consists of two interviews, one with the robot and the other with a human. During the intervention, the number of child's gazes towards the robot or human, the number of successful imitations, and the chain of repetitive and stereotyped behaviours were counted for evaluation. As a result of this intervention, the child's attention towards the human was slightly higher than towards the robot. The child's gaze towards the robot was higher than towards the human, and the child's imitation towards the human was slightly higher than towards the robot. The number of repetitive and stereotyped behaviours was significantly more while interacting with the human than with the robot. These results indicate that the child is more comfortable looking at the robot than the human during the interaction. Regarding the imitation analysis, the results indicate no significant differences in imitating either the human or robot. These findings show that the participants engaged with QTrobot, and it managed to keep them motivated, which resulted in a decrease in their repetitive behaviour.

3.2.2.2 Non-anthropomorphic Robot "Animal Robot"

Animal-assisted interventions are significantly increased due to the emotional, physical, and social relationships between animals and humans, especially those diagnosed with autism spectrum disorder (Smith and Dale, 2016). Wijker et al. (2017) suggests that the interaction between humans and animals improves social behaviour as they reduce the level of anxiety and stress.

According to Stanton et al. (2008), 22 children with ASD participated in a therapy session with a real dog for 15 weeks, show improvement in their language and social skills. In another study, children with ASD interact more with the therapists in the presence of a dog. They were happy, looking at the dog and interacting with it (Martin and Farnum, 2002). Despite the benefits of animals in medical therapy, the animal-assisted activity and interventions are limited due to the high cost of training animals to meet the required needs, safety, allergy, and ethical issues of using them (Sefidgar et al., 2016; Bharatharaj et al., 2017). Due to these limitations, researchers consider alternatives using robotic animals, which have demonstrated some improvement in mental health and have assisted therapists in enhancing the social skills of children with ASD.

Stanton et al. (2008) conducted a study to compare the social interaction skills of children with ASD between the robotic dog (AIBO) and mechanical dog (Kasha) as shown in Figures 7 and 8. Kasha can walk and make noises but does not respond to any interaction. 11 children with ASD (aged 5-8 years) participated in this study, including 10 boys and 1 girl. The children interact

Figure 7: AIBO.
(Stanton et al., 2008)

Figure 8: Kasha.
(Stanton et al., 2008)

with AIBO and Kasha, and then the time each child spent with each intervention, the amount of speech each child produced to each dog, and the behavioural interactions of each child during the interaction with each dog are measured. This study shows that children interacted and spoke more with the robotic dog AIBO than the mechanical one Kasha. This finding is in line with the previous studies that suggest robotic animals motivate children with ASD and make them engage in social interactions with humans. However, AIBO and Kasha are different in form and functions, and it would be better to compare two platforms with the same functions but with different appearances.

Bharatharaj et al. (2017) developed a parrot robot (KiliRo) to improve learning and social interaction skills and reduce stress. 24 children with ASD (aged 6-16 years) participated in this study that involved 49 sessions. The sessions were 60 minutes long and ran daily for 49 days. KiliRo could recognise voices and objects and was equipped with sensors that could recognise tactile interactions. As a result of this experiment, the robot helped increase the participants' social interaction skills and assisted in learning tasks. Additionally, there was a significant decrease in participants' stress levels after interacting with KiliRo. The generalisation of the acquired skills was not assessed.

A study by Panagiotidi et al. (2019) was employed the MiRo robot as a therapeutic tool for children with ASD. As shown in Figure 9, MiRo is a pet-sized mobile platform with six sensors and eight degrees of freedom. It can mimic animals' behaviours and responses, providing the psychological and emotional benefits of a pet. MiRo was designed to enhance human-animal interaction rather than human-human interaction. MiRo is expected to be used as a social care robot, educational robot, and research platform. 40 children with ASD (aged 11-16 years) took part in this study. Using MiRo as a therapeutic tool is still under investigation as the study is still underway.

Figure 9: MiRo Robot.
(Panagiotidi et al., 2019)

Another study by Ghafurian et al. (2020) utilised the MiRo robot to understand how people perceive emotions expressed by an animal robot. MiRo robot was programmed to express 11 affective expressions using colours for its LEDs and body postures. For example: happy = orange, sad = light brown, excited = red, disgust = green, surprise = white, annoyed = blue, tired = purple, bored = gray, angry = red, fear = pale blue or gray, and calm = no colour. 120 typically developing participants enrolled in this study (aged 23-69 years). The results show that half of the participants recognised happy, sad, excited, surprised and tired, but they could not recognise annoyed, angry and bored. Fear was recognised as a surprise by the majority of the participants, and boredom was recognised as tiredness. Some emotions were hard to recognise due to the lack of facial expressions of MiRo. Additionally, two emotions had the same colour (excited and angry), which was confusing. Several studies agreed that red represents anger. MiRo's features and movements also affected the accuracy of expression. It can only move forward, backward, left, and right with some head, ear, and neck movements.

Di Nuovo et al. (2020) conducted a study that utilised MiRo and NAO robots to evaluate their effectiveness in supporting children with ASD. 5 children with ASD (aged <5 years) participated in this study. The participants are expected to engage and interact with NAO and MiRo. The study is ongoing, but the initial results indicate that the children interacted more with NAO and neglected MiRo. The more advanced features of NAO (such as talking, singing, dancing) could explain this finding.

3.2.2.3 Critique of Research Outcomes

There has been substantial growth in the use of social robots for training the social skills of children with ASD. Children with ASD are more engaging in interaction with robots than humans (Aryania et al., 2020; Kumazaki et al., 2020).

The results from the conducted studies by Costa et al. (2018); So et al. (2019); Peca et al. (2020); Huijnen et al. (2021) revealed that the number of repetitive behaviours of children with ASD reduced while interacting with robots, and children also indicated more eye gaze and had better performance in imitation tasks in the interaction with robot compared to the human. This indicates that social robots make the sessions more fun and excited.

The conducted studies were employed different social robots with different characteristics, such as NAO, ZENO, QTrobot, KASPAR, MiRo, and AIBO. NAO was the most used robot among the reviewed intervention studies. 9 studies of the 24 studies included in this chapter were utilised NAO to train social skills and behaviours of children with ASD, including imitation, using gestures, turn-taking, joint attention, eye contact, play skills, and emotion recognition and expression skills. However, NAO does not have facial expression, it was employed in several studies (Miskam et al., 2014; Marinoiu et al., 2018; Shen et al., 2015) to train and enhance the emotion recognition and expression skills of children with ASD. NAO was programmed to express emotions using body gestures, colours, and audio cues. The findings indicate that NAO positively affects the participants' engagement and helps them recognise and express emotions. To highlight this, a study conducted by Cohen et al. (2011) to compare between NAO and iCat robot in expressing emotions. NAO expresses emotions using body gestures and colours, and iCat uses facial expressions as shown in Figure 10. The results show no differences between recognising emotions from iCat and NAO except for the sad emotion iCat was better.

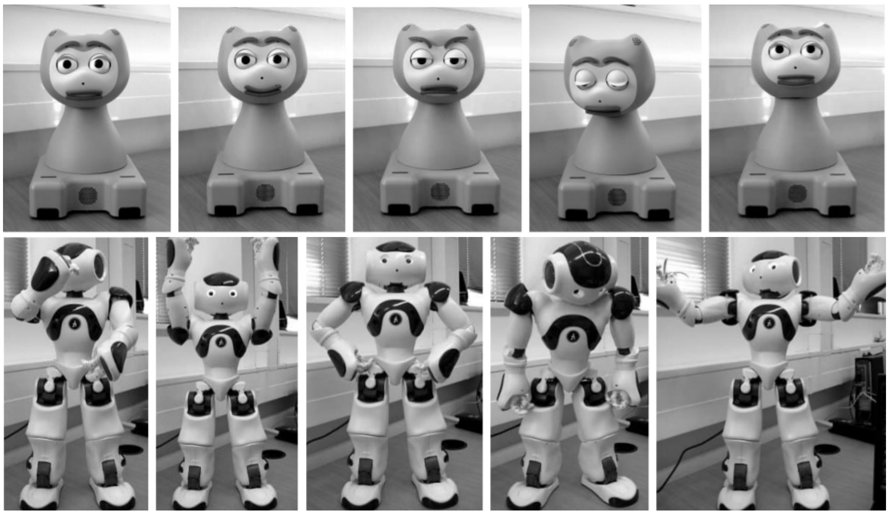


Figure 10: iCat and Nao express emotions: fear, happy, angry, sad, and surprise. (Cohen et al., 2011)

Children with ASD are more engaging with humanoid robots than animal robots (Kumazaki et al., 2020; Di Nuovo et al., 2020; Lee et al., 2012). The variety of functions that humanoid robots have, including dancing, singing, and talking, results in better efficiency in catching attention. To highlight this, Children with ASD interacted more with NAO and ignored MiRo in a study conducted by (Di Nuovo et al., 2020). These findings indicate that humanoid robots are better at training social skills such as joint attention, imitation, emotion recognition, and turn-taking, while animal robots reduce stress, anxiety, and repetitive behaviours. Additionally, the humanoid robots may facilitate generalising skills as they imitate some human features or behaviour.

The extensive review of robot-assisted interventions for children with ASD emphasised some limitations. The availability of robot-assisted interventions is minimal. The high cost of robot platforms is a significant reason for this limitation (Hung et al., 2019). For example, a fully-featured NAO robot costs $9000 dollars (RobotLab2020), and Zeno (Milo) robot costs $6500 and $3500 per year for the updated software (DELLTechnologies, 2020). Therefore, many families can not afford therapy sessions that incorporating a social robot.

Additionally, the efficient use of social robots requires being monitored by a technician or professional. Alcorn et al. (2019); Hughes-Roberts and Brown (2015) interviewed educators to understand the barriers to deploy social robots in real-life settings such as school or home. Some educators raise concerns about the need for experts to set-up the robot and emphasise that teachers with little technical knowledge will found the robot challenging to operate. Sigacheva et al. (2020); Diehl et al. (2012) emphasises the need for a trained or technical specialist to deploy social robots at home. Furthermore, it was observed the lack of involvement of parents and teachers in the intervention sessions (Diehl et al., 2012). The robot-based intervention has been focused mainly on the English language. Arabic is still in its early stages.

3.3 Research Design Methods

The research design refers to the overall strategy used to conduct research that defines the plan to answer the research questions by collecting and analysing data. Khowaja et al. (2020) categorised the research design into four types. The categories are post-test only, post-test control group, pre-test and post-test, and pre-test, and post-test control group.

- Post-test only: involves evaluating independent variables (such as length and type of the intervention) at the end of an intervention.
- Post-test control group: involves the evaluation of different groups at the end of an intervention.
- Pre-test and Post-test: involves evaluating dependant variables (such as social skills) before and after an intervention to measure the intervention's effectiveness.
- Pre-test and Post-test control group: involves evaluating variables in different groups before and after the intervention to measure the intervention's effectiveness on each group.

Most of the studies reviewed in this chapter used pre and post research method, revealing its significance in evaluating the interventions. This method helps to investigate the effectiveness of interventions by comparing the data collected in the pre-test with those in the post-test. The comparison helps indicate the changes or improvement in the participants' skills and behaviour after receiving an intervention.

3.4 Social Skills Assessments for Children With ASD

Social skills deficits are a hallmark characteristic of children with ASD. Several studies were conducted to evaluate the effectiveness of social skills interventions for children with ASD. These studies aimed to collect data to assess the generalisation and maintenance effects of the interventions (Rotheram-Fuller et al., 2013). The generalisation is the occurrence of behaviours or skills under conditions that differ from the conditions in which these behaviours or skills were targeted (Gunning et al., 2019). Maintenance means the child will continue to demonstrate the acquired skills over time, even after the training is finished (Rotheram-Fuller et al., 2013; Gunning et al., 2019).

Collecting data can be obtained through screening measures. Screening measures are psychological assessment methods used to evaluate children with Autism Spectrum Disorder (ASD), Dyslexia, Attention deficit disorder (ADD), Attention Deficit Hyperactivity Disorder (ADHD), Hyperlexia, and Sensory Processing Disorder (SPD). The most popular screening methods are interviews, observations, rating scales, self-report scales, and standardised tests (Rotheram-Fuller et al., 2013; Khowaja et al., 2020).

- Interviews: Allow the screener to gain insights from experts or users regarding an application's design, usability, and effectiveness. The experts included in the interviews were therapists, teachers, and parents to understand their behaviour's perspective. User interviews aim to understand the usability, feasibility, and effectiveness of the intervention.
- Observations: The screener observes the child's interaction with others, tools or technology interventions. The observation includes any gathered information during live or video observation of the child. Data can be collected either through coding scales, such as the Social Interaction Code and the Behaviour Coding Scheme (coding systems with known psychometric properties used to record specific social behaviours) or through observations of the child's skills and behaviours (such as gestures, initiations, responses, and affect) (Wendy Stone et al., 2010; Rotheram-Fuller et al., 2013).
- Rating Scales: Allow the screener to assess the child's skills and behaviours, such as the Social Responsiveness Scale (SRS), Social Skills Rating System (SSRS), Social Skill Improvement System (SSIS). The screener can be a parent (Parent-Rating) or a teacher (Teacher-Rating).

  Social Skills Rating System contains descriptors of social behaviours as affective understanding, joint attention, perspective taking, initiating interactions, and maintaining interactions. Social Responsiveness Scale is a quantitative measure to assess the level of impairment in social behaviours and to distinguish ASD conditions from other psychiatric conditions (Cunningham, 2012; Rotheram-Fuller et al., 2013).
- Self-Report Scales: Let the children perceive their skills and behaviours.

- Standardised Tests: These tests aim to evaluate how the child is achieving compared to other peers of the same age, such as Wechsler Intelligence Scales (uses to determine the child's intellectual ability "IQ"), Comprehensive Assessment of Spoken Language, Social Communication questionnaire (SCQ), Autism Spectrum Screening questionnaire (ASSQ), and NEPSY-II.

After reviewing the studies in Table 1, observations (Live and Video), questionnaires, and rating scales (Parent and Teacher reporting) are the most common methods to assess the social skills of children with ASD.

Table 1: Technology intervention studies for enhancing the social skills of ASD

Citation	**Target Skill**	**Participants: N,Age,Gender**	**Technology Used**	**Data Collection Methods**	**Setting**	**Findings and Outcomes**	**Generalisation**
Alcorn et al. (2011)	Follow eye gaze and understand clues	32 ASD, 5-14 years, 29 boys and 3 girls	Desktop VE	Video observation	Controlled research environment	Did not report any changes in the participants skills	Not assessed.
Abirached et al. (2011)	Emotional skills	9 ASD, 4-11 years, 7 boys and 2 girls	Desktop VE	Interviews and video observation	child's home or cafe	Did not report any findings regarding the participants skills.	Not assessed
Jarrold et al. (2013)	Social attention	37 HFA, 8-16 years, NM	Immersive VE	Standardised questionnaires (ASSQ, Wechsler,SCQ) and rating scales (SRS)	Not mentioned	Social attention of children with HFA is less compared to typically developing children.	Not assessed.
Lorenzo et al. (2013)	Social skills	20 HFA, 8-15 years, 16 boys and 4 girls	Immersive VE	Teacher interviews and observation	Not mentioned	Improvement of social skills.	Some skills transferred to the real classroom.
Bekele et al. (2013)	Monitor eye gaze and physiological differences	10 HFA and 10 TD, 13-17 years, 16 boys and 4 girls	Desktop VE	Eye tracking and physiological features	University lab	Differences between HFA and TD in recognising and processing faces	Not assessed
Serret et al. (2014)	Emotional skills	33 ASD, 6-17 years, 31 boys and 2 girls	Desktop VE	Observation	Controlled research environment	Improvement in the emotion recognition skills	Potential for generalisation.

Citation	Target Skill	Participants: N,Age,Gender	Technology Used	Data Collection Methods	Setting	Findings and Outcomes	Generalisation
Greczek et al. (2014)	Imitation skill	12 HFA, 7-10 years, 8 boys and 4 girls	NAO	Kinect	Controlled research environment	Improvement in the participants' imitation skills. Providing feedback managed to keep the participants motivated.	Not assessed.
Miskam et al. (2014)	Emotion expression	1 TD, 6 years, boy	NAO	Author designed questionnaire, video observation	Controlled research environment	Participant engaged with NAO	Not assessed.
Costa et al. (2014)	Emotion recognition	3 HFA, 12-15 years	Zeno	Video/live observation	Controlled research environment	Improvement in the participants' emotion recognition and imitation.	Not assessed.
Salvador et al. (2015)	Emotion recognition	11 HFA and 11 TD, 7-13 years, 15 boys and 7 girls	Zeno	Live/video observation	University campus	No significant differences between ASD and TD in recognising facial expressions. Significant improvement found in recognising fear and disgust after adding gestures.	Not assessed.

Citation	Target Skill	Participants: N,Age,Gender	Technology Used	Data Collection Methods	Setting	Findings and Outcomes	Generalisation
Cheng et al. (2015)	Interaction skills	3 ASD, 10-13 years, boys	Immersive VE	Video/live observation, standardised questionnaire (Wechsler)	Special school	Improvement between baseline and maintenance.	Not assessed.
Costa et al. (2015)	Identify body parts	8 ASD, 6-10 years, boys	KASPAR	Interview, author designed questionnaire, video/live observation	Classroom environment	participants leaned to identify body parts.	Participants generalised some skills to their classroom.
Chen et al. (2015)	Emotional skills	3 ASD , 10 -13 years, 2 boys and 1 girl	AR	Live observation	Controlled research environment	Improvement in the emotion recognition skills.	Not assessed.
Didehbani et al. (2016)	Social skills	30 HFA, 7-16 years, 26 boys and 4 girls	Desktop VE	Standardised questionnaires (NEPSY-II)	University of Texas	Improvement in emotion recognition skills, but not in social attribution.	Not assessed.
Lorenzo et al. (2016)	Emotional skills	40 ASD, 7-12 years, 29 boys and 11 girls	Immersive VE	Teacher interviews and observing facial expressions	Not mentioned	Improvement in emotional ability	Children generalised the learnt skills to real classroom.

Citation	Target Skill	Participants: N,Age,Gender	Technology Used	Data Collection Methods	Setting	Findings and Outcomes	Generalisation
Malinverni et al. (2017)	Social interaction	10 HFA, 4-6 years, NM	Desktop VE	Live and video observation	Specialised hospital	Reduce of repetitive behaviour.	Not assessed.
Bharatharaj et al. (2017)	Assessment of physiological changes	24 ASD, 6-16 years, NM	KiliRo	Observation and physiological measures	Specialised school	Significant decrease in the stress level. Helped increase the social interaction.	Not assessed.
So et al. (2018b)	Intransitive gesture	30 ASD, 4-6 years, 27 boys and 3 girls	NAO	Observation, standardised questionnaire (Wechsler) and author designed questionnaire	Specialised centres	Improvement in the gestures recognition and expression	Children generalised the learnt skills to novel setting.
Costa et al. (2018)	Attention and imitation skills	15 ASD, 4-14 years, NM	QTrobot	Rating scale (SRS), standardised questionnaire (Wechsler) and live observation	Controlled research environment	Child's gaze towards the robot was higher than towards the human. Significantly reduction in the repetitive behaviours while interacting with Qtrobot.	Not assessed.
Lee et al. (2018)	Greeting skill	3 ASD, 7-9 years, 2 boys and 1 girl	AR	Interview, standardised questionnaire (Wechsler), and observation	Treatment centre	Significant improvement in the participants' greeting skills	Not assessed.

Citation	Target Skill	Participants: N,Age,Gender	Technology Used	Data Collection Methods	Setting	Findings and Outcomes	Generalisation
Feng et al. (2018)	Imitation skill	2 ASD, 4 and 6 years, 1 boy and 1 girl	NAO	Live/ video observation	Not mentioned	Imitated the robot hand movements.	Not assessed.
Zorcec et al. (2018)	Greeting, imitation, and emotion skills	2 ASD, NM, 1 boy and 1 girl	KASPAR	Video/live observation	Hospital	Improvement in greeting skills. Participants managed to differentiate between Happy and Sad.	Generalised the taught skills.
Marinoiu et al. (2018)	Emotion recognition	14 ASD, 4-8 years	NAO	Video observation	Controlled research environment	Improvement in emotion recognition skills.	Not assessed.
Cao et al. (2019)	Imitation, turn taking, and joint attention	38 ASD, 3-5 years, NM	NAO	Programmatically, and live observation	Controlled research environment	WoZ achieved better results in turn taking tasks. Similar results in imitation and joint attention tasks.	Not assessed.
Moon and Ke (2019)	Communication skills	15 HFA, 10-14 years, 13 boys and 2 girls	Desktop VE	Video observation	Not mentioned	Emphasised the importance of visual stimuli. Difficulty in initiating interactions.	Not assessed.
Lorenzo et al. (2019)	Social communication	11 ASD, ages $<$ 6, 10 boys and 1 girl	AR	Author designed questionnaire	Controlled research environment	No significant differences.	Not assessed.

Citation	Target Skill	Participants: N,Age,Gender	Technology Used	Data Collection Methods	Setting	Findings and Outcomes	Generalisation
Herrero and Lorenzo (2019)	Emotion recognition	14 ASD, 4-15 years, 13 boys and 1 girl	Immersive VE	Author designed questionnaire and observation	Controlled research environment	significant improvement in the participants' emotion recognition skill	Not assessed
Wood et al. (2019)	Visual perspective taking	12 ASD, 6-14 years, 7 boys and 5 girls	KASPAR	Video Observation	Specialised school	Participants engaged in most of the activities.	7 children generalised the skills.
Stewart Rosenfield et al. (2019)	Conversation skill	1 ASD and 1 ADHD, 6 and 7 years, 1 boy and 1 girl	Immersive VE	Live/video observation	University research lab	Positive feedback from participants	Not assessed.
Soares et al. (2019)	Emotion recognition and expression	45 HFA, 5-10 years, 36 boys and 9 girls	Zeno	Live/video observation	Schools and centres	Significant improve in recognising and expressing emotions compared to other groups. 30% increase in the recognising emotions from social stories.	Generalised the gained skills.
So et al. (2019)	Gestures and eye gaze	23 LFA, 6-12 years, 20 boys and 3 girls	NAO	Video observation	Classroom	More eye contact toward the robots. No significant difference in the imitation skill between the robot and human sessions.	Generalised the taught gestures to new contexts.

Citation	Target Skill	Participants: N,Age,Gender	Technology Used	Data Collection Methods	Setting	Findings and Outcomes	Generalisation
Charlton et al. (2020)	Conversation skills	5 ASD, 8-10 years, 4 boys and 1 girl	Desktop VE	Author designed questionnaire, rating scale (SSIS), live and video observation	University Campus	Improvement in initiating conversation skills	Generalisation of the taught skills was observed.
Ke et al. (2020)	Communication skill	7 ASD, 10-14 years, 6 boys and 1 girl	Desktop VE	Author designed questionnaire, standardised questionnaire (SCQ) and video observation	Home	Improvement in the communication skills.	Not assessed.
Peca et al. (2020)	Imitation skill	4 ASD, NM, NM	NAO	Kinect and video observation	Not mentioned.	NAO has positive effect on participants. Capturing their attention.	Not assessed
So et al. (2020)	Joint attention and play skills	23 HFA, 4-6 years, 20 boys and 3 girls	NAO	Video observation	Specialised centres	Improvement in joint attention and play skills.	Generalised the skills with a human.
Huijnen et al. (2021)	Make a contact	9 ASD, 8-12 years, 8 boys and 1 girl	KASPAR	Video observation and interviews	Specialised school	Engaged more in the sessions with KASPAR than the teacher. KASPAR has positive effective on children with ASD.	Not assessed

3.5 Conclusion

Over the past two decades, the number of children diagnosed with or at risk for ASD has increased (Boat and Wu, 2015). The annual cost for diagnosis and treatment of ASD has been estimated at $236-262 billion annually: approximately $175-196 billion for adult services combined with $61-66 billion for children (Society, 2016). To reduce these costs, researchers emphasise the importance of assistive technology for training children and individuals with ASD (Desideri et al., 2020; Carter et al., 2016). This chapter focused on assistive technology for cognition. Assistive technology for cognition is any technology used to assist individuals with cognitive impairments to improve their daily performance and community participation. Virtual Environments (VEs) and social robots can be used as assistive technologies for cognition to enhance the social skills of children with ASD.

VEs are simulated platforms utilising 3D reality that the user can access via a screen or through more immersive technologies. Virtual environments are categorised into two types, immersive virtual environment and desktop (non-immersive) virtual environment. VEs have been used to enhance the social skills of children with ASD, as they offer a safe and enjoyable environment that eliminates the social anxiety of those children. Moreover, the animations in such environments are in line with research showing that children with ASD prefer programmes that include animation and sounds.

Social robots are also being increasingly used as an interactive and training tool for children with ASD. Evidence-based research has shown that children with ASD are more comfortable interacting with a robot than a human (Cao et al., 2020; Dickstein-Fischer et al., 2018), as robots offer consistent interaction and predictability (Alcorn et al., 2019; Diehl et al., 2012). These are precisely the sort of interaction that children with ASD prefer (Alcorn et al., 2019).

Although the promising results for using immersive virtual environments and social robots for training and enhancing the social skills of children with ASD, their widespread use is limited. The limited availability of immersive virtual environments and social robot interventions is due to the high cost of the platforms and the equipment needed. Moreover, the setup and control of such interventions require technicians or professionals as well as a dedicated environment. Despite the importance of generalisation in transferring the learnt skills from training interventions to real-life, most VE intervention studies do not conduct follow-up sessions to assess whether the children transferred the target behaviours and skills learnt to their daily lives. Similarly, the test of maintenance remains an open question in VE interventions.

Building on the knowledge gathered from the literature review of VEs and social robots interventions in training social skills of children with ASD; this study aims to develop an environment (Jammo-VRobot environment) that addresses some limitations of both interventions. The Jammo-VRobot environment will provide a hybrid approach of combining VE and a social robot to examine how their combined use would affect the social skills of children with ASD.

Non-immersive (desktop) VE is a cost-effective platform that promotes the availability of virtual learning environments and social robots. Additionally, a human-like embodiment is beneficial to promote the generalisation of the skill learnt through child-robot interaction to human-human interaction and engenders more engagement than the animal-like embodiment (Cao et al., 2020; Di Nuovo et al., 2020). Thus, the Jammo-VRobot environment is a desktop virtual environment that employs a virtual humanoid robot (Jammo-VRobot) to interact with children using an adapted social skill training programme. The social skill training programme for training children with ASD is adapted and modified from several studies (Huijnen et al., 2017; Costa et al., 2014; So et al., 2018b; Miskam et al., 2014; Soares et al., 2019). The availability of the Jammo-VRobot environment will make it easier for parents, teachers, and practitioners to use either at home or school. Using such a tool at home or school may improve the learning as it provides an additional resource for training the social skills. Additionally, it will ensure their involvement in the intervention and follow-up sessions. Lastly, the designed training programme will be guided by a parent or teacher to avoid concerns raised by some researchers and educators that such interventions may increase the social isolation of children with ASD. According to King et al. (2017), parents raise a concern that using iPad makes their children zoned in. Additionally, some educators state that some children are getting emotionally attached to the robot and starting isolating themselves, which results in detracting the interaction with their peers and family (Alcorn et al., 2019).

The Jammo-VRobot environment will benefit from the successful reviewed research studies. For example, the importance of social stories in training the emotion recognition and expression skills of children with ASD will be considered (Ghanouni et al., 2019; So et al., 2018b; Soares et al., 2019; Costa et al., 2014). Researchers emphasise that the reasons behind emotions should be explained to children with ASD (Costa et al., 2017; Ghanouni et al., 2019). Many studies indicate that children with ASD are visual learners, and they pay attention to visual cues (Soares et al., 2019; Costa et al., 2014). Therefore, the Jammo-VRobot environment will provide virtual environments to illustrate the context of social stories to facilitate the learning process. According to Charlton et al. (2020), using popular cartoon characters increases the children's attention and engagement towards an intervention. Thus, the Jammo-VRobot environment will use popular cartoon characters to motivate children and make the learning process fun and exciting. Greczek et al. (2014) observed that providing feedback keeps children with ASD motivated throughout the sessions. Additionally, Huijnen et al. (2017) discussed some essential requirements for social robots interventions. For example, the robot's sound should be computerised and friendly, and the sentences used should be simple and short. Moreover, therapists emphasise that the intervention sessions should not exceed 30 mins to match the attention span of children with ASD. These requirements will be taken into consideration while implementing the Jammo-VRobot environment.

4 Design of Jammo-VRobot environment

Building on the knowledge gathered from the current literature: on virtual environments, social robots, and their use in enhancing the social skills of children with ASD, an open-source tool that combines VE and a social robot is proposed to improve imitation skills, recognition and production of intransitive gestures and to train the emotion recognition skills of children with high-functioning ASD. The tool is a desktop virtual environment that employs a 3D virtual robot (Jammo VRobot) that interacts with the children through a training programme guided by a parent or teacher. The Jammo-VRobot environment aims to replicate the success achieved with physical robots to enhance the social skills of children with HFA. Implementing a desktop VE helps promote the availability of a virtual learning environment for children with ASD to enhance their social skills and provides a cost-effective platform for parents and practitioners.

4.1 Chapter Overview

In this chapter, the design considerations for developing a virtual environment for children with ASD are presented. This chapter focuses on the first two stages: pre-design and design. Design considerations from the most effective current interventions are surveyed, and then the design and development of the Jammo-VRobot environment are presented. Figure 11 shows the stages of designing and implementing a VE for children with ASD.

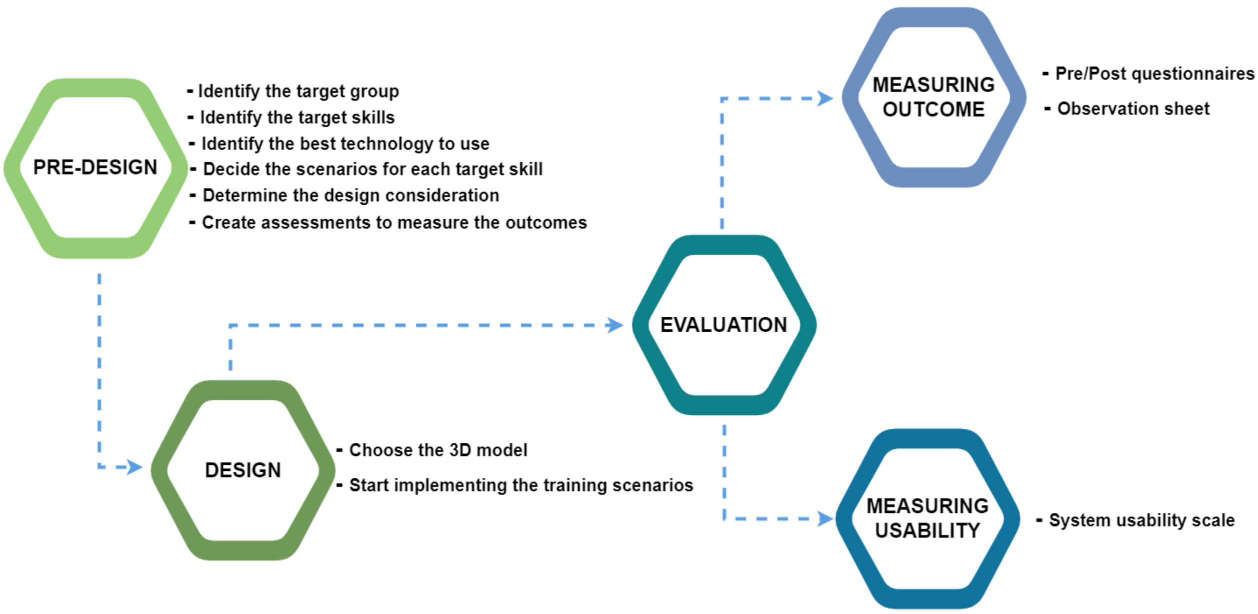


Figure 11: Diagram for the design and implementation of Jammo-VRobot environment.

Developing a successful tool for children with ASD highly depends on the pre-design and design phases. In these phases, different aspects of the tool should be considered such as the target group, learning objectives, the technology used, and the interaction design (Tsikinas and Xinogalos, 2020).

4.2 Target Group

The key to the success and acceptance of any tool or software is that it meets the needs of the intended users (Goh et al., 2008). A tool for children with ASD can be examined from different perspectives. For example, the specific condition such as high-functioning or low-functioning, the age group, and the purpose of the tool.

Jammo-VRobot environment aims to enhance the social skills of children with ASD; thus, it targets children with high-functioning ASD (HFA). Children with HFA are characterised by difficulties functioning in social situations, regardless of their higher cognitive and language abilities.

This research focuses on autism spectrum disorder as a cognitive disability. Therefore, the Jammo-VRobot environment requires children with no hearing, visual or physical impairments to interact with the tool. As early intervention can make a difference and produce positive and permanent changes in children with ASD (Cervera et al., 2011), the target age group is 4 - 12 years old.

The evaluation process was conducted in Egypt. Thus, the tool was developed in two languages: English and Arabic.

4.3 The Learning Objectives

The objective of the Jammo-VRobot environment is to enhance the social skills of children with ASD. Therapists and traditional training focus on turn-taking, emotion recognition and expression, imitation, joint attention, and the use of gestures. Therefore, this research targets three social skills: imitation, emotion recognition and expression, and intransitive gestures. These social skills were chosen for the following reasons.

4.3.1 Imitation Skill

Imitation skills are an essential element in the child's social learning process to develop new behaviour they need to mimic. Imitation is a prerequisite for verbal language as there is evidence that pre-verbal children use imitation as a communication tool (Stone et al., 1997; Ingersoll and Lalonde, 2010; Peca et al., 2020). Several psychiatric studies claim that children with ASD have a deficit in imitation (Ingersoll and Lalonde, 2010; Stone et al., 1997). This deficit includes object imitation (banging on a drum), gestures and body movement imitation, which causes serious implications for their social and communicative behaviours as eye contact, verbalisation, and language development (Ingersoll and Lalonde, 2010). Ingersoll and Schreibman (2006) and Ingersoll (2008) found that the efficiency of imitation of gestures and body movements predicts their language outcomes, and the imitation of objects is connected to the improvement of their play skill and the difficulty in imitating other children's behaviours affects their peer play. So, the development of imitation skills in children with ASD plays an essential role in acquiring joint attention skills.

Reciprocal imitation training (RIT) is a naturalistic behaviour intervention found to be effective in teaching object and gesture imitation in children with ASD during play, which yields to progress in their social, play, and language development (Ingersoll and Lalonde, 2010; Lowry, 2011).

4.3.2 Emotion Recognition Skill

Emotion recognition is an important contributor to social communication (Berggren et al., 2018; Manfredonia et al., 2019; Rice et al., 2015). Children with ASD have difficulties in recognising other's emotions. These difficulties include recognition of emotions from voice inflection, face, gestures, body language, and integrating multi-model emotional information in social context (Berggren et al., 2018). Improving the emotion recognition and expression skills of children with ASD can be extremely crucial for their development. A wide range of different interventions was developed to enhance and practice emotion recognition of children with ASD as (Costa et al., 2014; Miskam et al., 2014; Costa et al., 2017; Bekele et al., 2013; Serret et al., 2014; Marinoiu et al., 2018).

4.3.3 Intransitive Gesture Skill

Gestures are the movement of hands, head or other parts of the body and a form of non-verbal communication. Typically developing (TD) children use gestures to refer to objects before they produce verbal labels for these objects (Özçalışkan et al., 2017). Gestures are a prerequisite for the development of languages and an early sign of communication skills (Ökcün-Akçamuş et al., 2019; Özçalışkan et al., 2017). Children with ASD have delayed gesture comprehension and production development compared to their typically developing peers of the same age. These impairments result in a delay in language development in children with ASD (Ökcün-Akçamuş et al., 2019). According to previous studies, children with ASD have difficulties imitating meaningful gestures. According to So et al. (2018b,a), these challenges in imitation that children with ASD face are due to the difficulty they have in imitating humans, as they tend to feel a low level of interest towards people. As a result, technology interventions became an excellent alternative to teach such children gestures through imitation. Despite the excellent interaction between children with ASD and robots, a few studies were conducted to teach them how to recognise and produce gestures.

4.4 Technologies Involved

Various tools were reviewed and tested to create the 3D interactive environment, such as Second Life (SL), Open Simulator, and Unity 3D. For modelling and animation, Blender, Cinema 4D, Akeytsu software were used.

The aim was to use a game engine with the following main features: High-level programming, 3D support and Deployment on multiple platforms. The Unity 3D game engine was chosen for the following reasons: Unity engine can create 2D, 3D, virtual reality and augmented reality games, it is a multi-platform game development tool; the editor runs on Windows and Mac OS and can produce games for Windows, Xbox360, PlayStation, Mac, Wii, and Android platforms. One of its key features is the integration support with Blender, Maya, Cinema4D, and Akeytsu.

Unity also supports scripting in C#. For animating the 3D robot, Akeytsu and Mixamo were used.

To better comprehend the implementation details of the VE environment, an overview of how Unity 3D works follows.

4.4.1 Unity 3D

In Unity 3D, the game consists of a hierarchy of scenes, game objects and components. The Scenes can be game menus, individual levels, etc. They contain the game objects of the game. Each scene can have different game objects, decorations, and environments. Game Objects do not do anything on their own. On the other hand, components are a set of data that attach to the game object to change its appearance, behaviour, functionality and visibility. All game objects begin with a Transformation Component that defines their rotation, position, and scale in the scene. Scripting is an essential ingredient in all Unity applications. The Script allows the modification of the component properties and responds to the user input. It can only be executed when attached to an active scene's game object.

4.5 Interaction Design

Designing a tool for children with ASD requires specific design considerations. To create a tool design framework for children with high-functioning ASD, an extraction of design guidelines from existing tools and studies was conducted. According to the literature review, the design considerations for creating a tool for children with ASD are mainly divided into two categories: methods of information presentation and task design (Tsikinas and Xinogalos, 2020; Bozgeyikli, 2016; Bozgeyikli et al., 2018; Britto and Pizzolato, 2016; Pavlov, 2014).

4.5.1 Methods of Information Presentation

Design considerations regarding methods of information presentation are presented in five categories: sound, text, visual, animation, and tutorial.

4.5.1.1 Sound/Audio

As children with ASD are sensitive to sounds, researchers suggest avoiding the use of disturbing and explosive sounds as sirens or fireworks while developing an assistive tool for this group. Goals and tasks should be clear before playing each scenario and reinforced during the scenarios (Carlier et al., 2020). Thus, providing audio instructions and commands have been proven to provide better understanding (Bozgeyikli, 2016; Bozgeyikli et al., 2018; Britto and Pizzolato, 2016), but emphasises that the instructions need to be simple. Each sentence needs to be short as possible (Pavlov, 2014; Groba et al., 2018).

The Jammo-VRobot environment provides audio instructions for each scenario, explaining the task and goal to the child. For example, at the beginning of each scenario, the virtual robot "Jammo VRobot" explains the task to the child by saying "Now! we are going to play an emotion recognition game. I am going to express an emotion, and you are going to tell the name of this emotion. Let's start". During the scenario, the Jammo VRobot reinforces the goal of the scenario. For example, while the Jammo VRobot expresses each emotion, he asks the child "What is the name of this emotion?". All the instructions are simple and avoided using complex words.

4.5.1.2 Text

Text is an essential component of a User Interface (UI) (Pavlov, 2014). Text is used to annotate various interactive elements. Using simple textual language is essential while developing a tool for children with ASD. The text's font for button labels, menus and any written instructions needs to be clear as it helps to give the right amount of information needed. Due to the language processing difficulties in children with ASD, researchers advise avoiding the use of only written instructions, although it could be used as complementary to other methods (Britto and Pizzolato, 2016; Bozgeyikli, 2016; Bozgeyikli et al., 2018). Thus, written clues, besides other elements, increase the children's motivation to engage with the tool.

In the Jammo-VRobot environment, the text used for the button's label and in the bubble is clear and simple as shown in Figure 12. As bright colours have been proven to have a calming and soothing effect on children with ASD, white colour was chosen for the text, either in the buttons or the bubble. The tool used text as a complementary element to other methods, such as audio, animation, and pictures. For example, while the Jammo VRobot says and expresses an emotion, the name of the emotion appears in a bubble as a written clue, and a picture for this emotion also presents as shown in Figure 13.

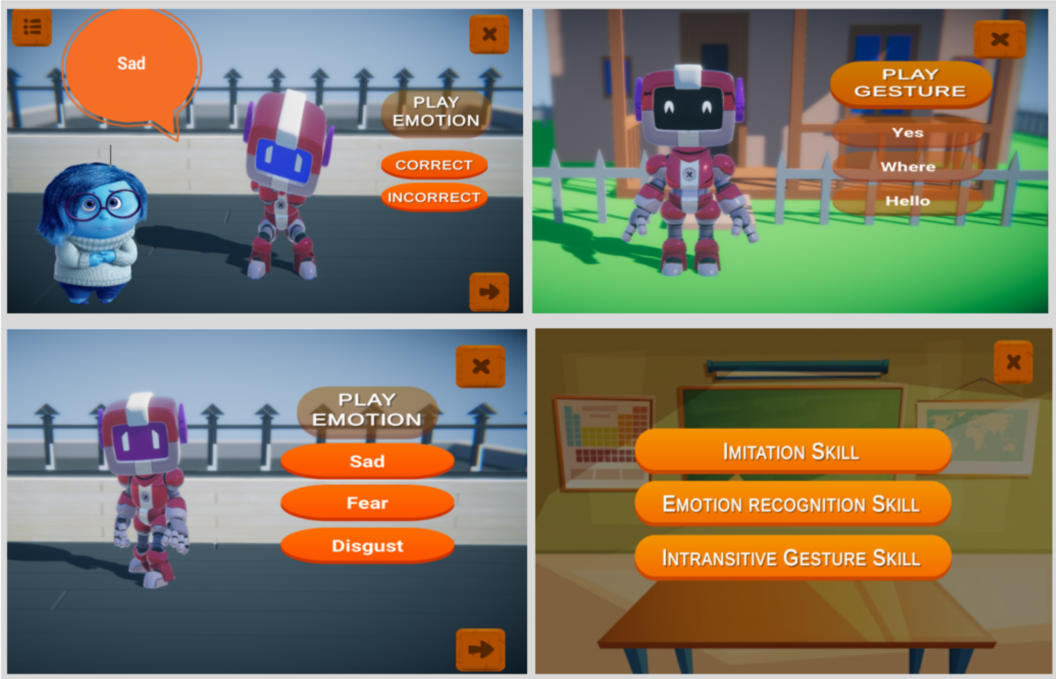


Figure 12: Examples of the text used for buttons and bubble in different scenarios in the Jammo-VRobot environment.

Figure 13: Text and picture as complementary elements besides the animation of the Jammo VRobot.

4.5.1.3 Visual

According to Britto and Pizzolato (2016); Kamaruzaman et al. (2016), colour should not be the only way to deliver content, but it has to be used as a supplementary element to emphasise information through other means. The contrast between the background and foreground is essential to distinguish items (Britto and Pizzolato, 2016).

Many studies emphasise that visual presentation may better grab the attention of children with

ASD. Children with ASD get attracted to large and colourful objects (Bozgeyikli, 2016). Buttons need to be big, stand out clearly from the other elements, provide appropriate click area, and ensure that they look clickable (Pavlov, 2014). Researchers emphasise the importance of having simplified graphics in facial expressions and having high-quality images (Bozgeyikli, 2016; Bozgeyikli et al., 2018). Designing a simple interface with few elements that presents only the features and content for the current task is advocated (Britto and Pizzolato, 2016). Navigation is another aspect that needs to be considered while developing a tool for children with ASD. Navigation needs to be simple and in a consistent location.

While developing the tool, colours were used as a supplementary element for expressing emotions besides the body postures, sound, change the shape of the eye, and images as shown in Figure 14. As the Jammo VRobot does not have facial expressions, colours were used to represent the basic six emotions: happiness, sadness, fear, anger, surprise, and disgust. Researchers have associated bright colours (e.g. pink, yellow) with positive emotions (e.g. surprise, happy) and dark colours (e.g. red, blue) to negative emotions (e.g. angry, sad) (Sutton and Altarriba, 2016; Ou et al., 2004). According to Sutton and Altarriba (2016) and Plutchick's emotion wheel, (happiness=yellow, disgust=green, anger=red, sadness=blue, fear=purple, surprise=pink). As children with ASD are visual learners (Xin and Sutman, 2011; Raisingchildren, 2018), each emotion was associated with a cartoon character Disney (2020), that was representing emotions to grab the child attention and help in the learning process as shown in Figure 14.

Figure 14: Colours and cartoon characters as supplementary elements for expressing emotions. (Disney, 2020)

Additionally, different virtual environments were designed for the storytelling scenarios, taking into account the plot of each story as shown in Figure 15, as a visual cue for the children to motivate them and increase their attention.

Figure 15: Different virtual environments for each story.

The navigation buttons are in a consistent location in all the scenes and use the standard icon buttons to avoid any confusion as shown in Figure 16.

Figure 16: Navigation buttons from different scenes.

Additionally, all the buttons are big enough and stand out clearly from other elements in the scene. Some features were added to the buttons to make them look clickable. At the beginning of each scenario, the choices buttons are disabled and lowered their opacity to make the focus on the play button and to be easily spotted. After pressing the play button, the choices buttons are enabled, and the opposite happened to the play button as shown in Figure 17.

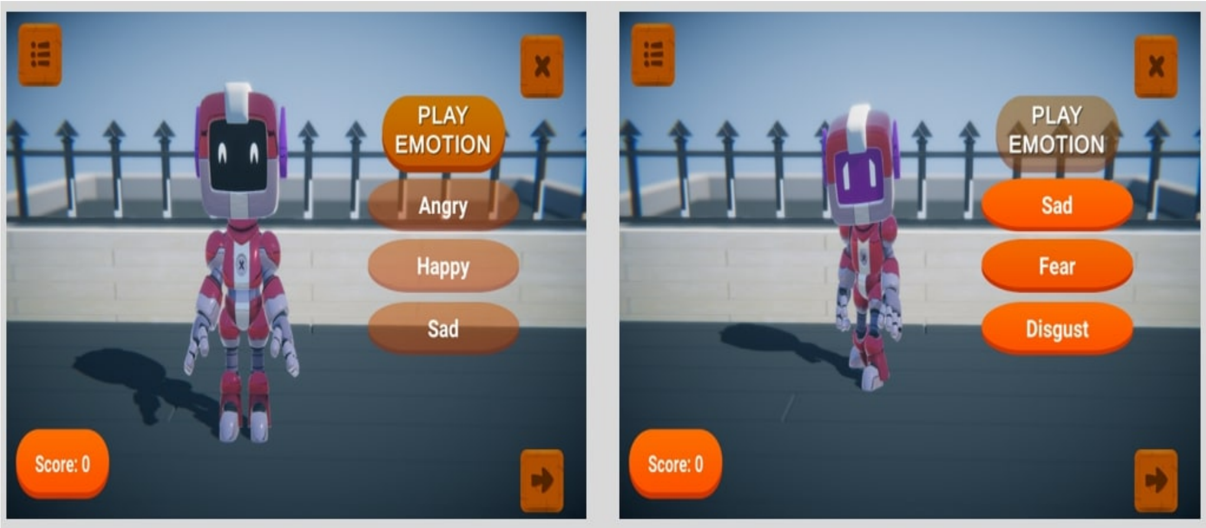


Figure 17: Button features.

4.5.1.4 Animation

Children with ASD are attracted and motivated to 3D animations. For example, animation helps the child to focus on the desired area (Britto and Pizzolato, 2016). Although too many animations may distract the child from the tasks. Thus, avoiding unnecessary animations were emphasised to aid focus (Bozgeyikli, 2016).

The Jammo-VRobot environment uses simple animations to provide the child with the information needed. The movement animation helps to attract the children's attention and makes them focus on the desired area. The Jammo VRobot uses pointing animation to draw attention to specific areas to help the learning process as shown in Figure 18. Additionally, the point animation was used as a supplementary element besides the sound. For example, the Jammo VRobot points at the bubble and says, "A picture of this emotion will appear here".

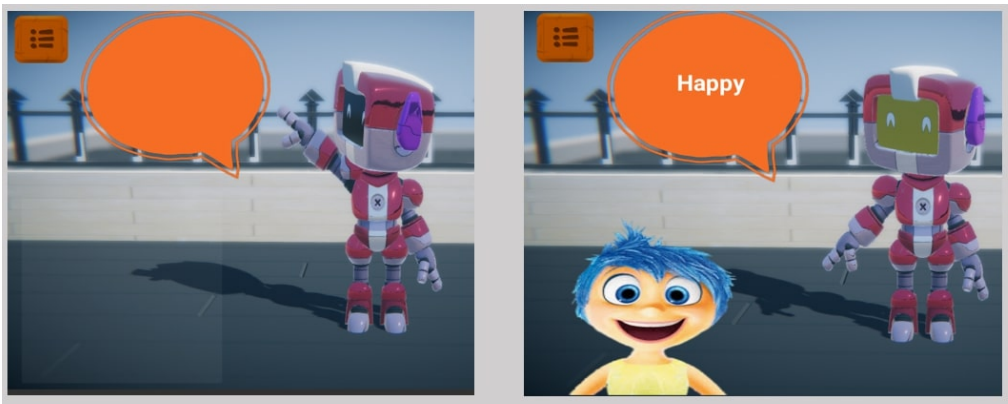


Figure 18: Point animation.

4.5.1.5 Tutorial

Since children with ASD show a low level of interest towards other humans, virtual teachers have been reported to work well with children with ASD. According to Bozgeyikli (2016); Bozgeyikli et al. (2018) using animated characters (e.g., dogs, robots or cartoon characters) rather than human characters can increase the child's motivation. Instead of employing virtual human characters to teach the children and give them feedback, using alternative characters (e.g. dogs, robots, etc.) was suggested as a better design principle for children with ASD (Britto and Pizzolato, 2016). Researchers emphasise that using animations and images while narrating a story or speech works well for tutorials and helps children with ASD to concentrate (Bozgeyikli, 2016). According to Howlin et al. (1999), using videos for tutorials is confusing for children with ASD.

4.5.2 Task Design

Design consideration regarding task design for children with ASD includes exploiting the ASD characteristics, task complexity, and feedback.

4.5.2.1 Exploiting the ASD Characteristics

Many studies suggest having a routine and a consistent structure in the training while developing a tool for children with ASD. These aspects meet the characteristics of children with ASD (Valencia et al., 2019; Bozgeyikli, 2016; Kamaruzaman et al., 2016). One of the essential design considerations for children with ASD is consistency throughout all the scenes, levels or scenarios. While training children with ASD, it is important to state clear goals and objectives to avoid any distractions. While the short memory skill of children with ASD, repetition of the tasks has been proven to provide better training for them (Bozgeyikli, 2016). To facilitate this, randomness in tasks was suggested since the users with ASD were observed to repeat the correct way over and over once learnt. Children with ASD are characterised by having a short attention span (Bozgeyikli, 2016), so the training sessions need to be short in order not to lose concentration.

The Jammo-VRobot environment generates the emotions and gestures randomly for better training to facilitate the repetition of the tasks. At the beginning of each scenario, the Jammo VRobot provides instructions and explains what will happen during the sessions to clarify each scenario's objective.

4.5.2.2 Task Complexity

Gradually increasing the task complexity during training has been observed to work better for children with ASD. In order not to overwhelm the child while interacting with the assistive tool, the intensity of the task elements as sound and crowdedness should be increased gradually following advice given in (Bozgeyikli, 2016). An important aspect regarding design consideration for task complexity is avoiding providing unachievable goals to avoid breaking the child's motivation.

All the goals in the proposed training programme are achievable, as evidenced in previously cited training programmes. The Jammo-VRobot environment does not provide any background sound to avoid any distraction for the children with ASD.

4.5.2.3 Feedback

Providing real-time feedback is a common usability recommendation while developing a tool for children with ASD (Britto and Pizzolato, 2016; Bozgeyikli, 2016; Valencia et al., 2019). Feedback is important to guide children while interacting with the tool and help them understand the application behaviour and predict the behaviour of similar elements or features. Giving positive feedback increases the child's motivation towards the task. Feedback can be given using animations, audio, text, and images (Bozgeyikli, 2016; Britto and Pizzolato, 2016). Negative feedback is inclined to make the child with ASD anxious (Bozgeyikli, 2016; Valencia et al., 2019). Thus encouraging words are advised to be used in any negative situation. Using a visible score as progress tracking with the tool are observed to provide comfort for children with ASD (Bozgeyikli, 2016). However, other studies suggest avoiding competitive elements as a score because it created a negative feeling for children with ASD (Bozgeyikli et al., 2018; Bozgeyikli, 2016).

The Jammo-VRobot environment provides real-time feedback regarding the child's action, as Jammo VRobot either compliments the child when the answer/action is correct using animation and audio by saying "very good" or encourages the child to try again if the answer/action is wrong using animation and audio by saying "let's try another one". As negative statements of failure make children with ASD anxious (Valencia et al., 2019; Bozgeyikli, 2016), the Jammo VRobot encourages the child when the answer or action is wrong by saying "let's try another one" instead of "wrong answer" or "bad job".

Table 2 summarises the design considerations that were considered while developing the tool for children with ASD.

Table 2: Design considerations for the Jammo-VRobot environment.

Item	Explanation	Category
Sound/Audio	The Jammo-VRobot environment provides audio instructions about what will happen in the session/scenario at the beginning to the child.	Methods of Information Presentation
	The given instructions are simple and avoided using complex words.	
Text	The text used for the button's label is clear and simple.	
	The Jammo-VRobot environment uses text as a complementary element to other methods. Ex: The name of the gesture appears on the screen along with the animation and sound.	
	Bright colours were used for the text to calm the children.	
Visual	The Jammo-VRobot environment does not use the colour as the only way to provide information.	
	The navigation button is in a consistent location at all the scenes to avoid confusing the child.	
	The location of the buttons used is clear and do not hide any information in the background.	
	The buttons used are big enough and stand out clearly from the other elements in the scene.	
	Standard icon buttons were used to avoid any confusion.	
	Different virtual environments were designed as a visual cues.	
	Cartoon characters were associated with the text, sound, and colour.	

Item	Explanation	Category
Animation	The Jammo-VRobot environment uses simple animation to provide the child with the information needed.	Methods of Information Presentation
	Animations were used to make the children focus on specific areas.	
Tutorial	The virtual robot was employed as a tutor.	
Exploiting the ASD Characteristics	The elements (e.g. buttons) are consistent in all scenes.	Task Design
	The Jammo-VRobot environment provides generating the emotions and gestures randomly to facilitate the repetition of the tasks.	
Task complexity	All the set goals in the training programme are achievable.	
	To avoid overwhelming the child through the training, no background sounds provided.	
Feedback	The Jammo-VRobot environment provides real-time feedback regarding the child's action, as the robot either compliment the child when the answer/ action is correct using animation and audio (Very Good) or encourage the child to try again if the answer/action is wrong using animation and audio (Let's try again/Let's try another one).	
	Avoid using negative statements for failure.	

4.6 Virtual Robot Characteristics and Animation

Before creating the virtual environment, the 3D robot model as shown in Figure 19, was chosen. The 3D model "Jammo Character" was chosen because of its humanoid robot-like appearance as studies emphasise that humanoid-like embodiment might help children with ASD generalise the taught skills (Cao et al., 2020; Di Nuovo et al., 2020). The Jammo VRobot is the 3D model in the Unity Asset Store (UnityAssetStore, 2020). It has a large number of degrees of freedom to present upper and lower body gestures. This research is built on the success of physical robots (NAO, Zeno, QTrobot) in training the social skills of children with ASD. Each of these robots has different characteristics and appearance; thus, the Jammo VRobot was chosen to keep consistency in all the training scenarios and avoid confusing the child.

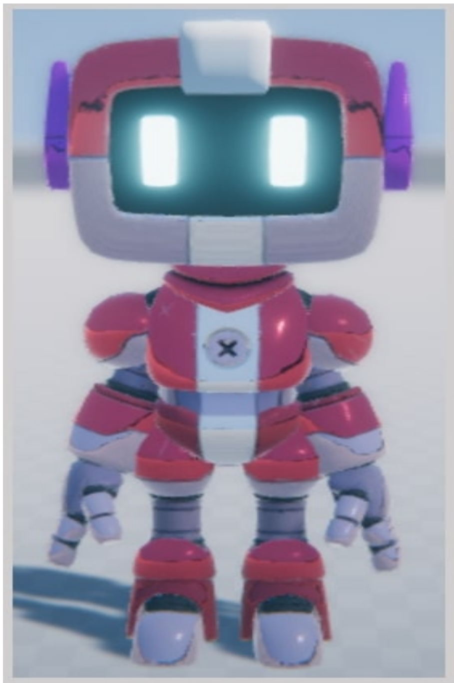


Figure 19: The Jammo VRobot.

For Jammo's voice, several audio clips were recorded using Text-to-Speech (TTS) online software (Wideo, 2020). Audacity software, audio editor software, was used to change the intonation of the sound to make it sound like a child (Audacity, 2020). Then, all the audio clips were imported into Unity. Once imported into Unity 3D, the animations and audio clips were assigned to the Jammo VRobot via scripts and audio components. Each component is related via code to some tag to be activated by the moderator/observer during the interaction. The appropriate scripts were attached to trigger the animations. During each scenario, the user can interact with the environment by clicking the relevant button.

4.6.1 Generation of the 3D Animations

3D animation combines different motions applied to a 3D object (characters, vehicles, etc.) to bring it to life. Different systems are available for creating 3D animations such as 3ds Max, Blender, Mixamo, Cinema 4D, Clara.io, iclone, Maya, and Akeytsu.

For generating the 3D animations for Jammo VRobot, Mixamo and Akeytsu were used. Mixamo (2020), is an online platform for creating 3D animations. It has an extensive library of animations and characters. In this free software, users upload their 3D characters and mark some key elements like knees and elbows. Then, they can choose from the existing animations on the website and assign them to the 3D model as shown in Figure 20. Additionally, it is possible to download a ready animated 3D model from the website as it contains various 3D characters. After applying the needed animations to the 3D model (Jammo VRobot), the animations were exported using the Unity supported format "FBX" to be imported into Unity.

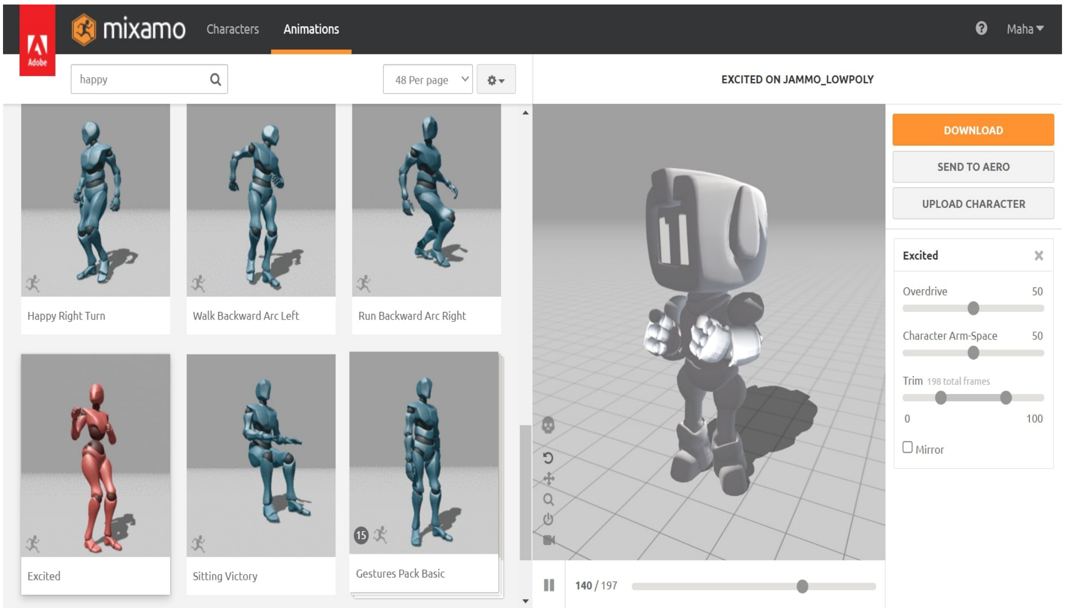


Figure 20: Exporting animations from Mixamo: the animations on the left can be applied to the Jammo VRobot. For example, exciting animation was applied to Jammo

Although Mixamo provides thousands of animations, not all the needed animations for the training programme were available. Thus, Akeytsu was used to create these animations. Akeytsu is a key-frame animation software that helps to create any animation from scratch as shown in Figure 21. It allows for importing and exporting FBX files; thus, it is easy to integrate the created animation into Unity 3D.

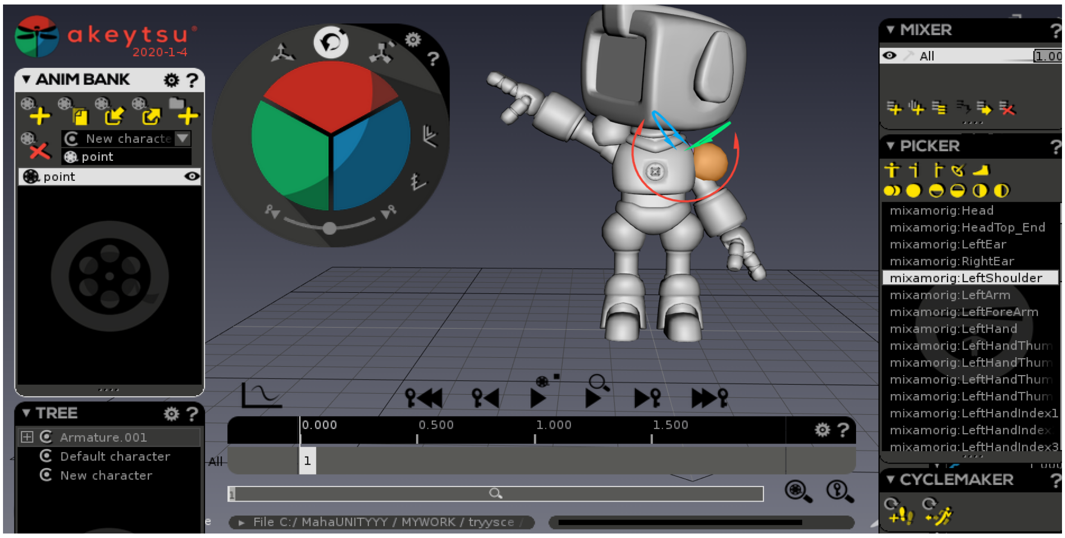


Figure 21: Akeytsu software for animation. Animated the pointing animation.

4.6.1.1 Emotional Gesture Development

The focus of the research in this thesis is on the basic six emotions by Ekman (1999), which includes happiness, fear, sadness, anger, surprise, and disgust. A set of gestures was developed to represent these emotions. The emotional gestures were created by looking at the commonalities in head and body positions in expressions as portrayed by non-ASD children in the studies were conducted by (English et al., 2017; Calvo et al., 2015) and collected in the robotic emotional gesture database used by (So et al., 2018b,a). These emotional gestures were created using Mixamo and Akeytsu and applied to the Jammo VRobot as shown in Figure 22. A description of each emotional gesture can be found in Table 3 along with the intended emotion.

Table 3: Description of the emotional gestures.

Emotion	Form of gesture
Happy	Jammo VRobot sways from side to side with excitement move of its arms
Sad	Jammo VRobot bends its knees and keeps bringing its arms toward its face
Angry	Jammo VRobot has both arms crossed then both arms on its hip and turns his head abruptly
Surprise	Jammo VRobot places both hands toward its face
Disgust	Jammo VRobot turns its head and places its hand in front to the head as a rejection
Fear	Jammo VRobot moves backwards

Figure 22: Emotional gestures expressed by the Jammo VRobot.

4.6.1.2 Intransitive Gesture Development

Eleven gestures commonly used in daily life were chosen for this social skill training programme. The social gestures children are being trained to recognise and use are: hello, where, come, yes, not allowed, good job, look at this, awesome, hungry, stop, and me. These gestures were designed by looking at cards with PECS (Picture Exchange Communication System) as shown in Figure 23, representing gestures used by specialists for teaching children with ASD Miller-Wilson (2006), and gestures performed by NAO robot in previous studies (So et al., 2018b,a).

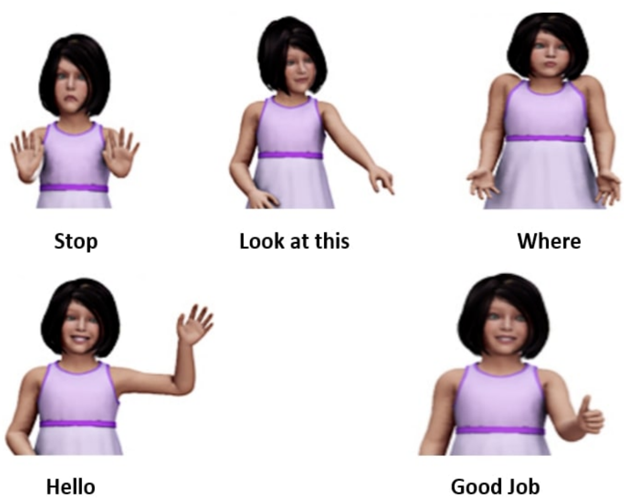


Figure 23: PECS cards represent gestures.
Miller-Wilson (2006)

The 11 gestures were designed using Akeytsu software and integrated into Unity 3D as shown in Figure 24. A description of each intransitive gesture illustrates in Table 4.

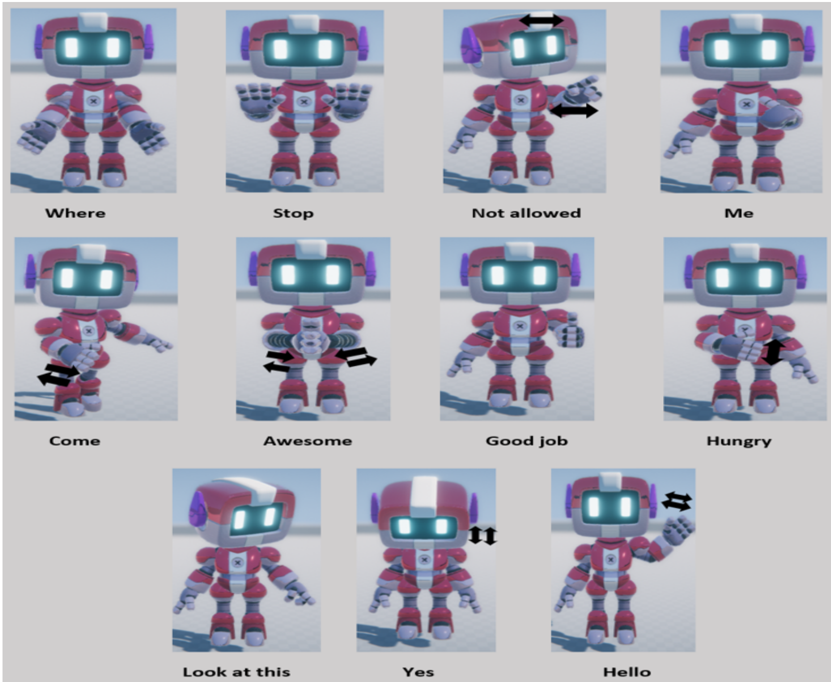


Figure 24: The intransitive gestures expressed by the Jammo VRobot.

Table 4: List of the 11 gestures producing by Jammo VRobot.

Gesture	Form of gesture
Hello	Right/Left palm waving
Yes	Head nods up and down
Good Job	Hold thumb finger up at the chest level
Stop	Right and Left hands hold at the chest level with the palm facing outward
Look at this	Index finger point at something
Where	Both arms wide open and face upward
Me	Index finger point to the chest
Awesome	Clapping both hands
Come	Right/Left arm extend and right/left palm crook and move towards self
Hungry	Right/Left hand move up and down at the lower chest level
Not allowed	Head nods right and left and index finger waving

4.7 Conclusion

As previously outlined in this chapter, designing computer-based training applications for children with ASD requires specific guidelines to promote a successful learning environment. This chapter presents the core principles of developing VE technologies that previous studies shared for children with ASD. These principles are mainly divided into two categories: methods of information presentation and task design. The presented guidelines and design considerations not only covered the VE technologies but also covered computer-based interventions.

While developing the desktop virtual environment, the mentioned design considerations were considered ensuring that the Jammo-VRobot environment meets all the needs of children with ASD. These considerations include sound, text, animation, visuals, tutorial, feedback, exploiting the ASD characteristics, and task complexity. Additionally, different aspects of the tool were determined, such as the target group, technologies used, and the learning outcomes (objectives). Children with HFA aged between 4-12 years are the target user group. Unity 3D game engine, Mixamo, and Akeytsu were used to design and implement the Jammo-VRobot environment. The learning objectives of the Jammo-VRobot environment is enhancing the imitation, emotion recognition and expression, and intransitive gestures of children with HFA.

As the proposed social skill training programme replicates the success of physical robots (NAO, Zeno, QTrobot) with different characteristics and features, the decision was made to use one 3D model of a humanoid robot (Jammo VRobot) to maintain consistency in the training programme scenarios. Additionally, it was impossible to replicate the physical robot's appearance as these companies do not publicise the 3D model. Therefore, Jammo VRobot was chosen. The Jammo's animations were designed in Mixamo and Akeytsu then integrated into Unity.

The main idea behind the contribution of this research is to investigate whether a virtual robot (Jammo) could be as effective as physical robots in training the social skills of children with HFA. The social skills training programmes used by physical robots were improved by developing a desktop virtual environment (Jammo-VRobot) that combines these several training programmes in one unified training programme. The next chapter presents the structure of the proposed social skill training programme..

5 Development of the Training Programme

5.1 Chapter Overview

This chapter focuses on the social skills scenarios that were designed and developed for children with HFA. The proposed social skill training programme aims to train emotional ability, improve imitation, and enhance the intransitive gestures of children with HFA. The proposed social skill training programme encourages the triadic interaction between Jammo VRobot, parent/teacher, and child.

5.2 Technical Decisions

The software used in the development of the Jammo-VRobot environment includes:

- Unity 3D in its 2019.3 version as the game engine.
- Akeytsu and Mixamo for the animation of the humanoid robot (Jammo-VRobot) in the virtual learning environment.
- Audacity software to create the humanoid robot's voice.

The Unity 3D game engine was chosen as it allows for different programming languages, but the C# was used due to its compatibility with the program. Another reason for choosing the Unity 3D game engine is its integration support with several animation software such as Akeytsu, Maya, Blender, and Cinema4D.

The 3D model was chosen from the Unity Asset Store (UnityAssetStore, 2020). The Jammo character is a free 3D model with several textures. The second step was creating the needed animations for the virtual learning environment. Mixamo was the first platform used to create some animations. The fbx. file (3D character) was uploaded to the Mixamo platform to start applying the ready animations on it. After applying the needed animations to the Jammo 3D model, the animations were exported using the Unity supported format .fbx to be imported into unity. The second option for creating the rest of the animations was Akeytsu software. Akeytsu allows creating any animation or sequence of animations from scratch by moving and rotating the joints and different parts of the 3d model. The next step was designing and creating the user interface (UI).

Buttons are one of the most used UI components. First, a Canvas was created in the Hierarchy as a parent of all the UI components, including buttons and score as shown in Figure 25. The button has two components, the button itself and the text. The default text was replaced with the TextMesh Pro. The TextMesh Pro has more options than the default text. In the Inspector, the text font and colour were adjusted to meet the design considerations for children with ASD. On the button itself, the button component is the most important component as it controls the button's functionality. In the On Click () function, the object that will be affected by the button is placed. Several scripts were created and attached to that object (Jammo-VRobot) to be run when the button is clicked as shown in Figure 26.

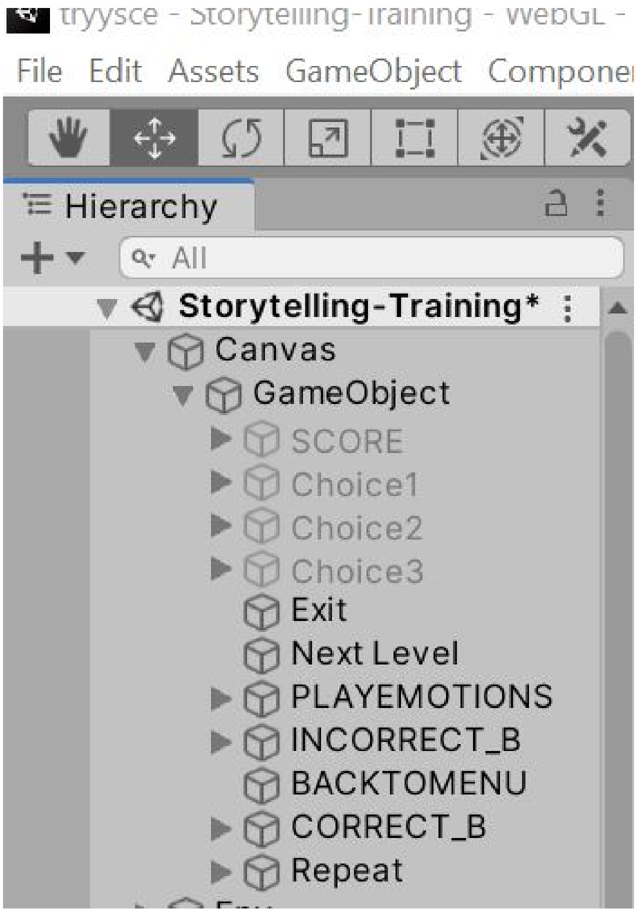


Figure 25: The Canvas parent, buttons, and score as it appears in the hierarchy.

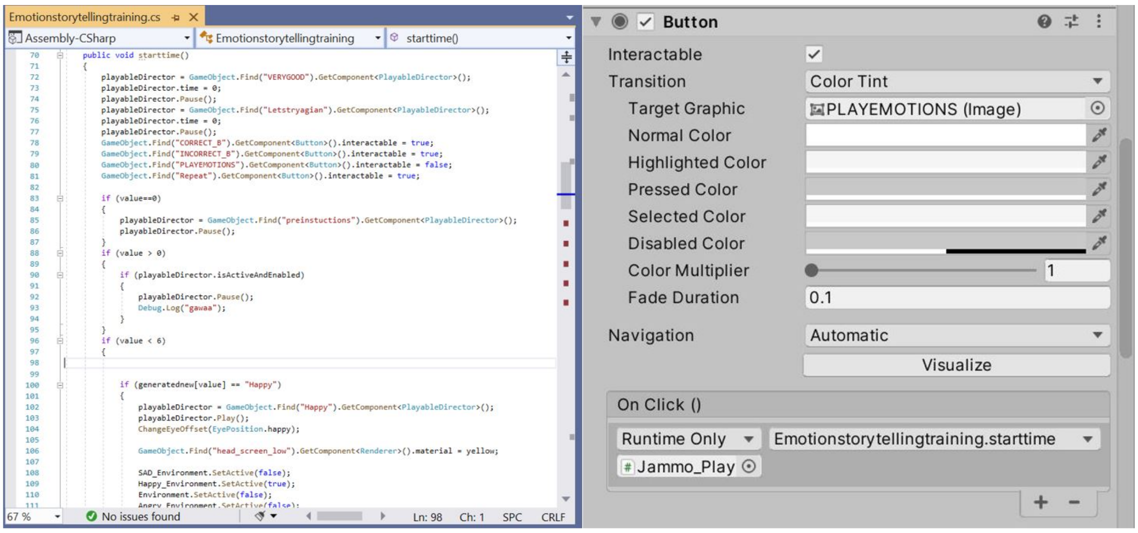


Figure 26: The button components, functions, and the a sample of script attached to the object

Figure 27 shows phase I of the intransitive gesture training programme and a sample of the code. The scene consists of the 'Play Gesture' button, 3 choices buttons, navigation buttons, and score. Once the moderator (teacher/parent) clicks the 'Play Gesture' button, the Jammo-VRobot animates one of the 11 gestures randomly, and three choices appear randomly on the options buttons.

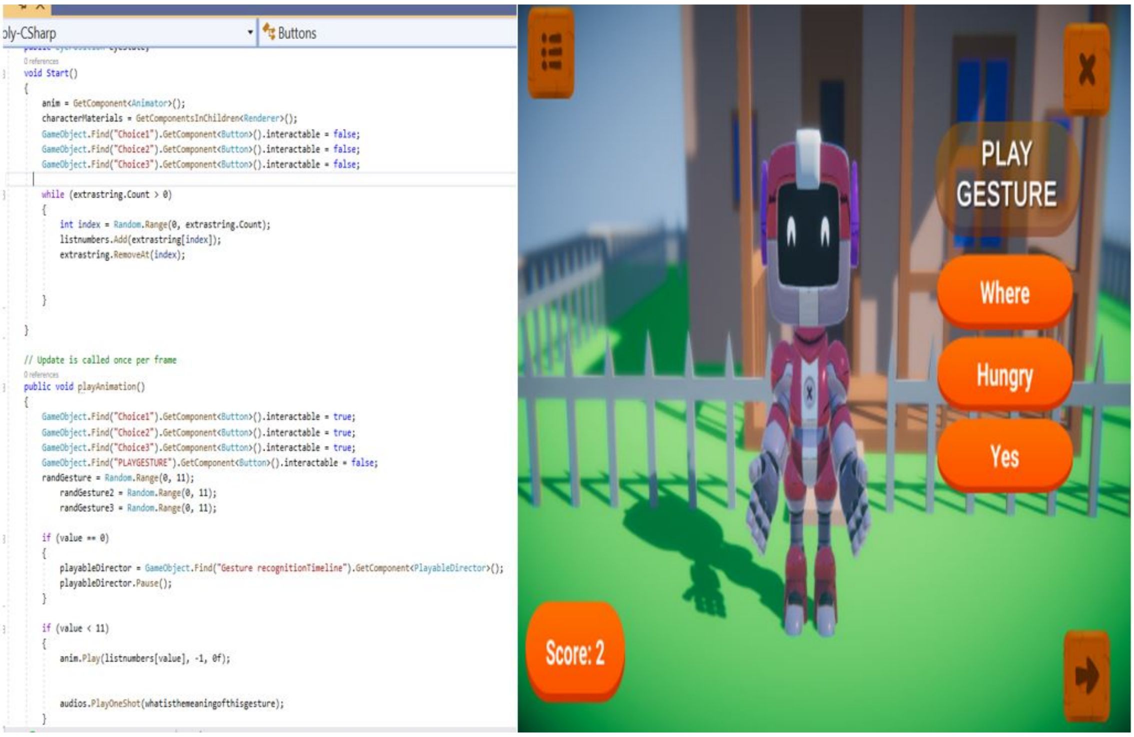


Figure 27: Fragment of code to activate the buttons.

Figure 28 shows the code for the scoring system. Each time the child answers correctly, the Jammo-VRobot compliments the child by saying "Very good" and performing motivational animation, and the score counts one. If the child's answer is wrong, the Jammo-VRobot encourages the child to try another.

```
public void whenchoose1()
{
    audios.Stop();
    selectedAnswer = R1;
    choiceSelected = "y";
    GameObject.Find("Choice1").GetComponent<Button>().interactable = false;
    GameObject.Find("Choice2").GetComponent<Button>().interactable = false;
    GameObject.Find("Choice3").GetComponent<Button>().interactable = false;
    GameObject.Find("PLAYGESTURE").GetComponent<Button>().interactable = true;

}
public void whenchoose2()
{

    audios.Stop();
    selectedAnswer = R2;
    choiceSelected = "y";
    GameObject.Find("Choice1").GetComponent<Button>().interactable = false;
    GameObject.Find("Choice2").GetComponent<Button>().interactable = false;
    GameObject.Find("Choice3").GetComponent<Button>().interactable = false;
    GameObject.Find("PLAYGESTURE").GetComponent<Button>().interactable = true;

}
public void whenchoose3()
{

    audios.Stop();
    selectedAnswer =R3;
   choiceSelected = "y";
    GameObject.Find("Choice1").GetComponent<Button>().interactable = false;
    GameObject.Find("Choice2").GetComponent<Button>().interactable = false;
    GameObject.Find("Choice3").GetComponent<Button>().interactable = false;
    GameObject.Find("PLAYGESTURE").GetComponent<Button>().interactable = true;

}
public void checktheanswer()
{
    while(choiceSelected=="y")
    {
        if(value-1 ==selectedAnswer)
        {

            anim.Play("Very Good");
            audios.PlayOneShot(Goodjob);
            count++;
            GameObject.Find("ADDSCORE").GetComponentInChildren<Text>().text = "Score: " + count ;
            break;

        }
        else
        {
```

Figure 28: Fragment of code to check the correct answers and counts the score

5.3 Design and Implementation of the Training Scenarios

The design of social scenarios was the first step towards developing the Jammo-VRobot environment. The review of the social skills deficits in children with HFA highlighted the importance of three main social skills, imitation, emotion recognition and expression, and intransitive gestures, and these are the focus of the training scenarios devised in this chapter. The interactive social scenarios were designed to enhance the imitation skill of children with HFA, train their emotion recognition and expression ability, and improve the recognition and production of intransitive gestures. The scenarios are delivered across a number of different sessions which are divided into three phases: greeting, training, and finishing session as shown in Figure 29. All the sessions start with Jammo VRobot greets the child by saying, "Hello, my name is Jammo VRobot", followed by him explaining the scenario and what is expected in the session. The Jammo VRobot guides the child throughout the session and encourages him/her to complete the task by saying "Now I am going to express an emotion, and you are going to say the name of this emotion", "What is the name of this emotion?". After completing the entire training session, the Jammo VRobot ends the interaction by saying, "Now we have finished. bye-bye". The event flow is outlined in Table 5.

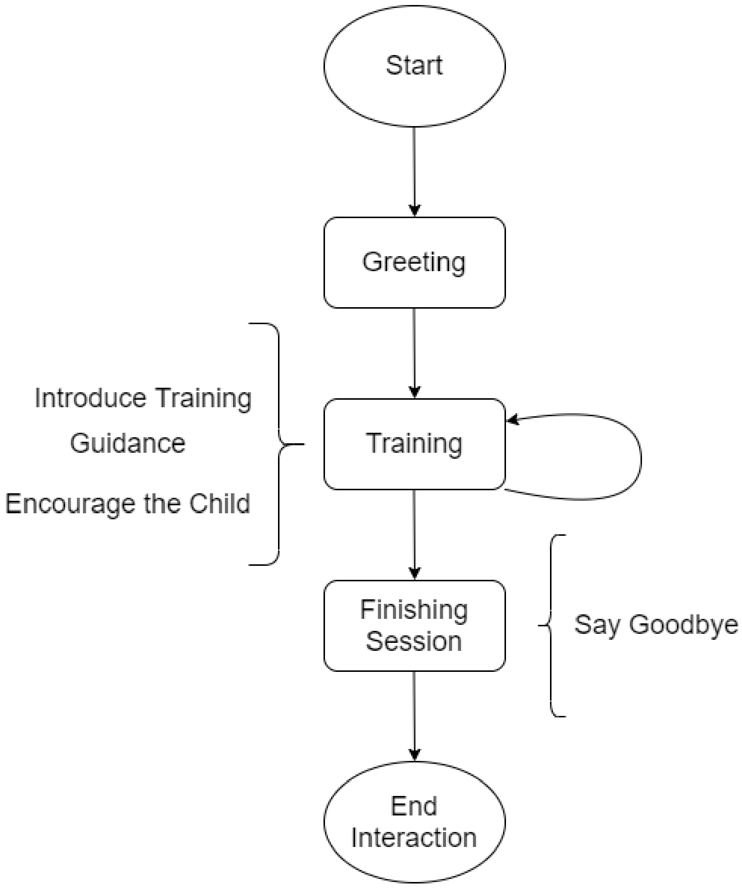


Figure 29: Phases of the interaction sessions

Table 5: Overview and explanation of the basic event.

Phase	**Basic event**	**Explanation**
Greeting	Greet the child	Jammo VRobot greets the child with the body movement and speech.
Training	Introduce training	Jammo VRobot explains the scenario.
	Guidance	Jammo VRobot guides the child to complete the task by language. For example, the robot says, "What is the expression for happy?".
	Encourage the child	When the child completes the task correctly, Jammo VRobot compliments him; when the child completes the task incorrectly, Jammo VRobot encourages him to try another one.
Finishing session	Say goodbye	Jammo VRobot says goodbye to the child by language and body movement.

The interaction scenarios build on several training programmes by (So et al., 2018b; Huijnen et al., 2017; Costa et al., 2014; Soares et al., 2019; Lux AI, 2019). The training programme involves three actors: a child, a parent/ teacher, and the Jammo VRobot. Each actor is assigned a specific role in the scenario. For example, an initiator (Jammo VRobot) who demonstrates the intended behaviour, a responder (child) who imitates the initiator's action or responds to the Jammo's questions, and an observer or a moderator (teacher or parent) who controls the movement of Jammo VRobot, observes the interaction between the child and Jammo VRobot and helps the child throughout the session if the child needs assistance. These three actors establish a triadic interaction, which happens between the child, the Jammo VRobot, and the teacher/ parent. Educators prefer the triadic interaction as they do not need the child to connect too much with the robot, as they might prefer to be with the robot rather than with parents, siblings, or friends (Alcorn et al., 2019).

The social skill training programme used in this intervention is informed by established work and existing practices that were done using physical robots and proven to be beneficial in enhancing the social skills of children with ASD (Huijnen et al., 2017; Costa et al., 2014; So et al., 2018b; Soares et al., 2019; Lux AI, 2019). Figure 31 shows the training programme's structure. The social skill training programme consists of three skills training programmes: imitation skill training programme, emotion recognition and expression training programme, and intransitive gesture training programme as shown in Figure 30.

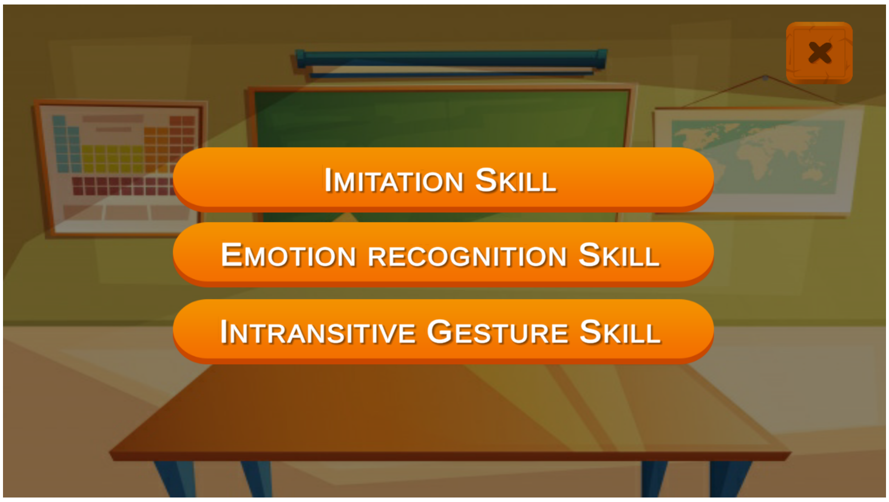


Figure 30: The social skill training programme.

The imitation skill training programme consists of one scenario that targets imitating hand gestures, while the other two skills training programmes are divided into three phases, each phase contains three scenarios. Figure 31 presents the structure of the whole social skill training programme.

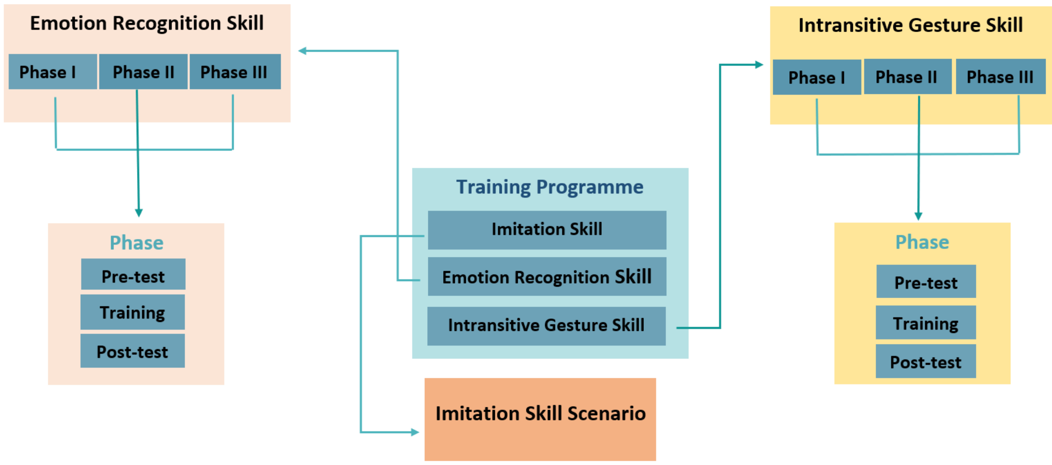


Figure 31: The structure of the social skill training programme structure.

5.3.1 Imitation-Based Interactive Scenarios

Children with ASD have difficulty with imitation (Ingersoll and Schreibman, 2006). Researchers emphasise that imitation skill plays a crucial role in developing the communication and social skills of children with ASD. Their ability to imitate hand gestures and body movements predict their language outcome, and their ability to imitate actions increase their joint attention (Ingersoll and Lalonde, 2010; Ingersoll, 2008).

Imitation is a good indication of the child's interest and engagement towards an intervention (Ingersoll, 2008). Therefore, the imitation skill training scenario is at the beginning of each intervention session. This scenario focuses on imitating hand gestures. The imitation scenario structure was adopted and modified from (Huijnen et al., 2017; Costa et al., 2014).

Jammo VRobot was programmed to animate nine different hand gestures: right arm up, left arm up, both arms up, right arm to the side, left arm to the side, both arms to the side, right arm to the front, left arm to the front, and both arms to the front as shown in Figure 32.

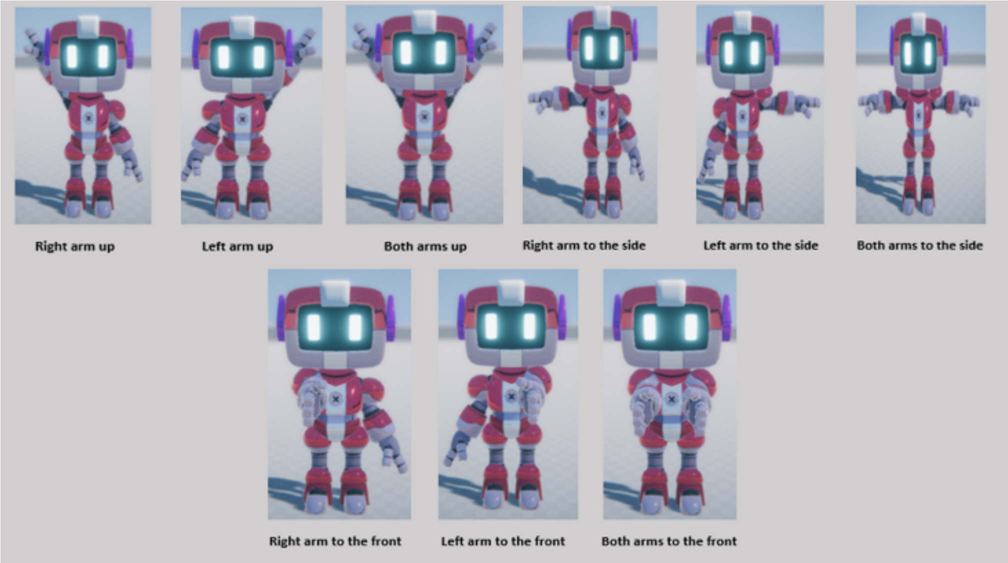


Figure 32: Jammo VRobot performs hand gestures.

After the Jammo VRobot finishes the instructions, the moderator/observer clicks the 'Hand Gesture' button to instruct the robot to animate a hand gesture. Jammo VRobot animates nine hand gestures one at a time in a randomised order while encouraging the child to imitate (e.g. "Both arms to the side"). The name of the motion appears in a bubble. The moderator/observer judges the child's action, and either clicks the 'Correct' or 'Incorrect' button. The Jammo VRobot gives feedback about the child's motion. For example, if the child makes the motion correctly, Jammo compliments the child by saying "Very good" and performs motivational animation. Whereas, if the child's motion is incorrect, the robot says, "Let us try another one" and performs a 'No' animation. Afterwards, the moderator/ observer clicks the 'Hand Gesture' button again to instruct the robot to animate another hand gesture. Figure 33 shows the hand gesture imitation scenario. The script of the scenario and the flowchart are shown in Figures 34 and 35 respectively.

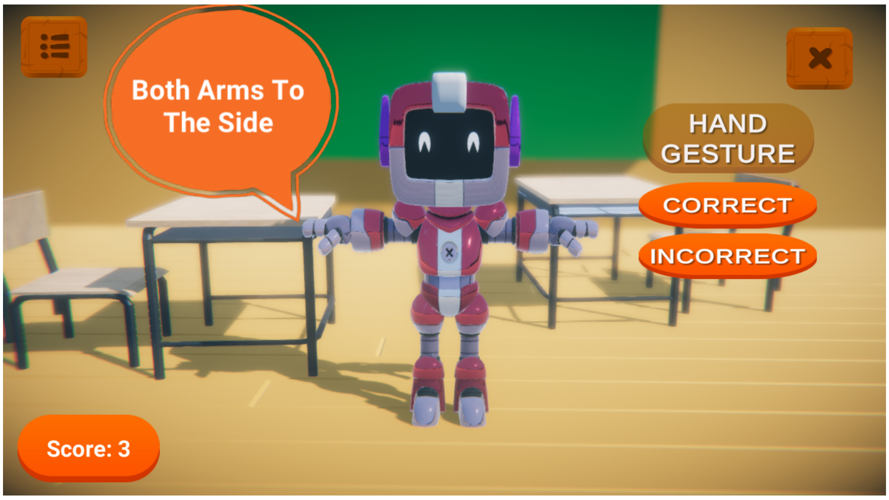


Figure 33: The hand gesture imitation Scenario.

Hello! My name is Jammo Robot.

[Performing "Hello" gesture]

Now! We are going to play an imitation game. I am going to do something, and you are going to do the same. Let's go.

[Performing "Talking" gesture]

The moderator presses " Hand Gesture" button to play one of the 9 hand gestures.

One arm up.

[Performing "Left hand up"- wait for the child's imitation]

If the child's motion is correct !

Very Good!

[Performing "Very Good" gesture]

If the child's motion is incorrect !

Let's try another one!

[Performing "No" gesture]

The moderator presses the "Hand Gesture" button

After finishing all the 9 hand gestures

Now! We have finished. Bye Bye

[Performing "Goodbye" gesture]

Figure 34: The hand gesture imitation script.

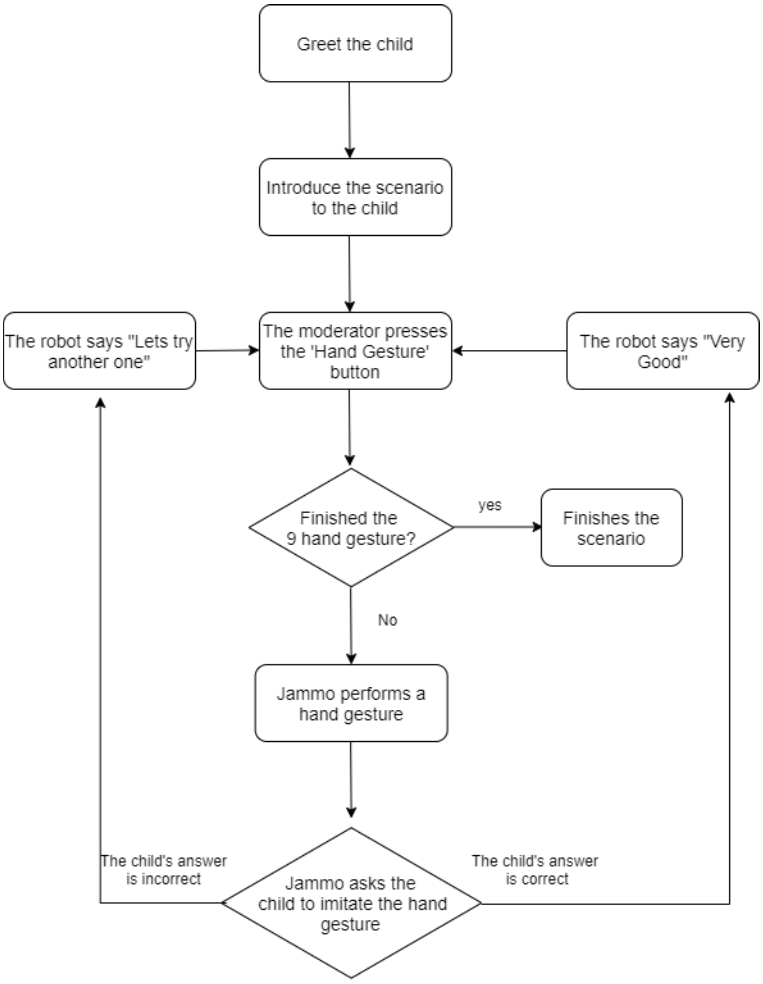


Figure 35: Flowchart of the hand gesture imitation scenario.

5.3.2 Emotion Recognition and expression Interactive Scenarios

The emotion recognition skill is a principal contributor to social communication skills. Improving the emotion recognition and expression skills of children with ASD is extremely crucial for their development. The emotion recognition training programme was built on emotion recognition teaching programme by Hadwin et al. (1996) that is widely used by the state-of-the-art research (Lux AI, 2019; Costa et al., 2017, 2018; Miskam et al., 2014; Soares et al., 2019).

Hadwin et al. (1996) developed a training programme to teach emotion understanding to children with ASD through 4 simple steps:

- Recognising basic emotions (e.g. fear, happiness, sadness).
- Understanding the causes of these emotions.
- Expressing the basic emotions.
- Understanding own and other's emotional state.

Children with ASD have difficulties recognising others' emotions, expressing emotions, and associating the relevant emotion to a social context. Thus, the emotion recognition training programme is divided into three phases, each addressing one of these difficulties. Each phase contains a pre-test, four training sessions, a post-test, and a follow-up post-test after two weeks. Phase I targets recognising and understanding the causes of these basic emotions. Phase II targets expressing the basic six emotions, and phase III targets understanding the affective state of others.

Jammo VRobot was programmed to express the six basic emotions introduced by Ekman (1999): Happiness, Sadness, Fear, Anger, Disgust, and Surprise. Jammo VRobot does not have facial expressions; therefore, body movements, colours, and eye shapes were used to express these emotions as shown in Figure 36. Table 6 describes the body movement of each emotion.

Figure 36: Jammo VRobot expresses six basic emotions.

Table 6: Body movement of each emotion.

Emotion	Form of gesture
Happy	Jammo sways from side to side with excitement move of its arms.
Sad	Jammo bends its knees and keeps bringing it arms toward its face.
Angry	Jammo has both arms crossed then both arms on its hip and turns his head abruptly.
Surprise	Jammo places both hands toward its face.
Disgust	Jammo turns its head and places its hand in front to the head as a rejection.
Fear	Jammo moves backwards.

5.3.2.1 Phase I

Phase I provides training for emotion recognition and focuses on the basic 6 emotions. Jammo VRobot greets the child and gives instructions for the pre-test. The purpose of the pre-test is to evaluate whether the child can recognise the six emotions expressed by the Jammo VRobot through body movements and changes in face colour and eye shape. The moderator/observer presses the 'Play emotion' button to instruct the robot to produce an emotion one at a time in a randomised order (e.g. both arms crossed, then both arms on its hip and turns his head abruptly). Each time Jammo VRobot asks the child, "What is the name of this emotion?" the child should choose one of the three options that appear on separate buttons on the screen as shown in Figure 37. The child either points at the correct answer (e.g. angry) or verbally says the name of the emotion. Jammo VRobot provides feedback based on the child's answer by either complimenting the child or encourages them to try another one. There is a score that counts the child's correct answers. The wrong answer does not reduce the score. The pre-test is completed when the six emotions have been covered. The flowchart and the scenario's script are shown in Figures 38 and 39.

Figure 37: Emotion recognition: Phase I-pretest.

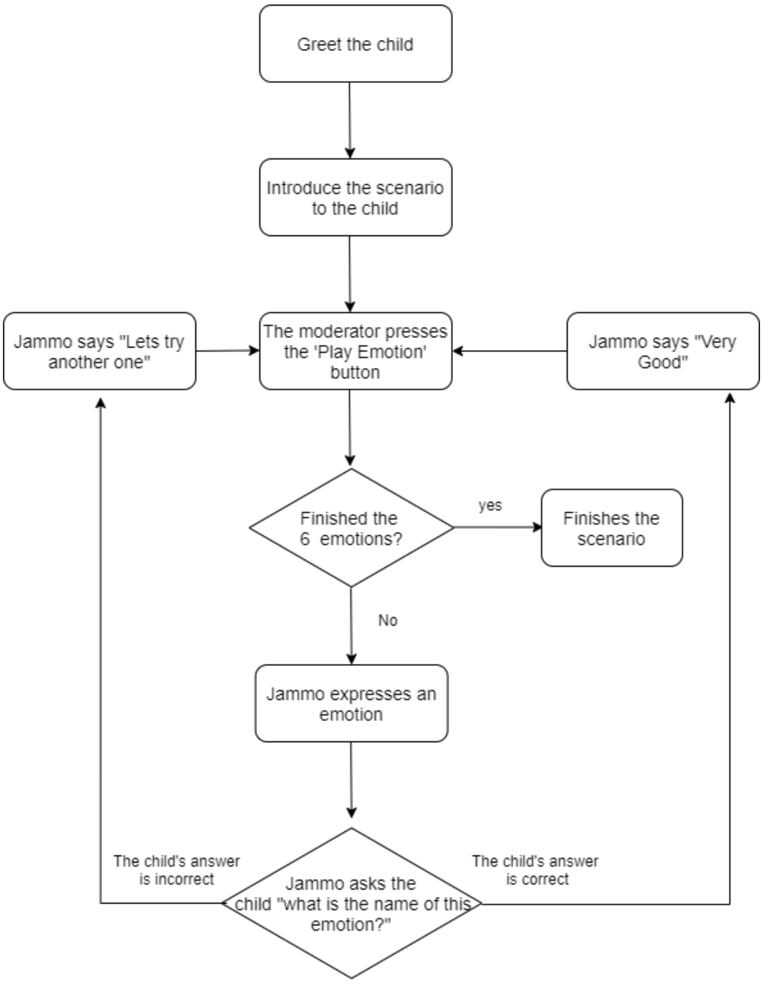


Figure 38: Flowchart of the emotion recognition: Phase I - pretest.

Hello! My name is Jammo Robot.
[Performing "Hello" gesture]
Now! We are going to play an emotion recognition game. I am going to express an emotion, and you are going to tell the name of this emotion. Let's go.
[Performing "Talking" gesture]

The moderator presses " Play Emotion" to play one of the six emotions.
What is the name of this emotion?
[Performing the emotion gesture, color and eye shape]
[Wait for the child's answer]

If the child's answer is correct !
Very Good!
[Performing "Very Good" gesture]

If the child's answer is incorrect !
Let's try another one!
[Performing "No" gesture]
The moderator presses the "Play Emotion"

After finishing all the six emotions
Now! We have finished. Bye Bye
[Performing "Goodbye" gesture]

Figure 39: Emotion recognition: phase I - pre-test script.

Once the pre-test is complete, the child proceeds to the training scenario. In the training scenario, the child watches the robot expresses the six emotions one at a time in a randomised order while indicating the emotion's name, when it is most likely felt, and highlighting the colour that representing this emotion. A cartoon character is associated to each of the emotions Disney (2020) to make learning emotions fun and engaging as shown in Figure 40. According to Charlton et al. (2020) using popular cartoon characters increases the engagement of children with ASD with the training interventions. The scenario starts with the Jammo VRobot greeting the child by saying "Now, we are working on emotions, the name of the emotion will appear here, and a picture of this emotion is here". Jammo VRobot then points at the cartoon character and says: "Look at this face", then names the emotion and says "This is a SAD face". Then Jammo starts describing when the emotion is most likely felt by saying "When something bad happens, we feel sad". Jammo then expresses the emotion using body movements, colour, and eye shape. The colour representing the emotion is also displayed on Jammo's face while saying "Blue is the colour for sadness". The name of the emotion "Sad" then appears in the bubble. Table 7 explains the basic six emotions. The same scenario is repeated for each of the 4 training sessions; however, the emotions are randomly generated each time.

Figure 40: Emotion recognition: Phase I-training.

Table 7: Explaining the basic six emotions.

Emotion	Description
Happy	When something good happens, we feel happy.
Sad	When something bad happens, we feel sad.
Angry	When someone bothers us, we may feel angry.
Surprise	When something happens that we do not expect, we feel surprised.
Fear	When something scares us, we feel fear.
Disgust	When we see or smell something nasty, we feel disgusted.

After the 4 training sessions are complete, the child then takes the post-test, which is identical to the pre-test. Two weeks after the training, the child is expected to take the follow-up test. The scores of the pre, post, and follow-up tests will be compared to analyse the training programme's effectiveness.

5.3.2.1.1 Phase II

Phase II aims at expressing emotions through imitation. The pre-test determines the child's ability to express emotions. Jammo VRobot asks the child to express an emotion one at a time in a randomised order by saying "What is the expression for happy?". The word 'happy' appears in a bubble as shown in Figure 41. The moderator/observer judges the accuracy of the expressed emotion and either presses the 'Correct' or 'Incorrect' button. The running score is then adjusted to reflect the answer. The flowchart and the scenario's script are shown in Figures 42 and 43.

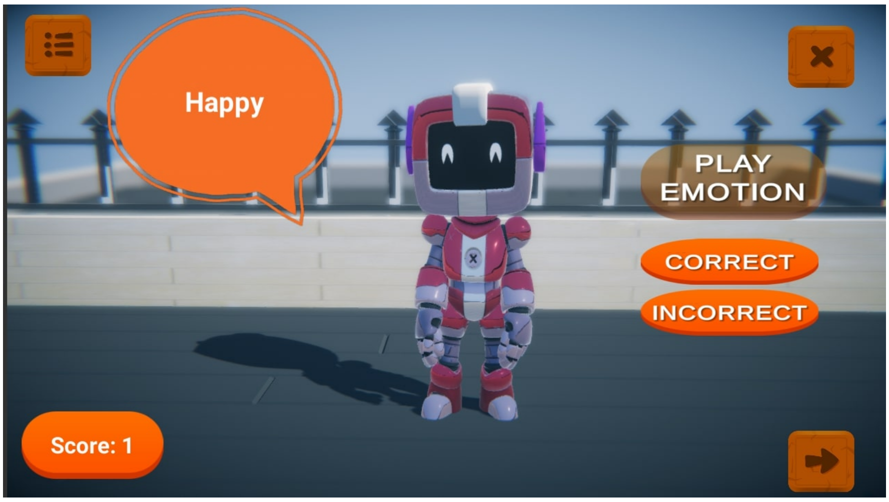


Figure 41: Emotion recognition: Phase II - pretest.

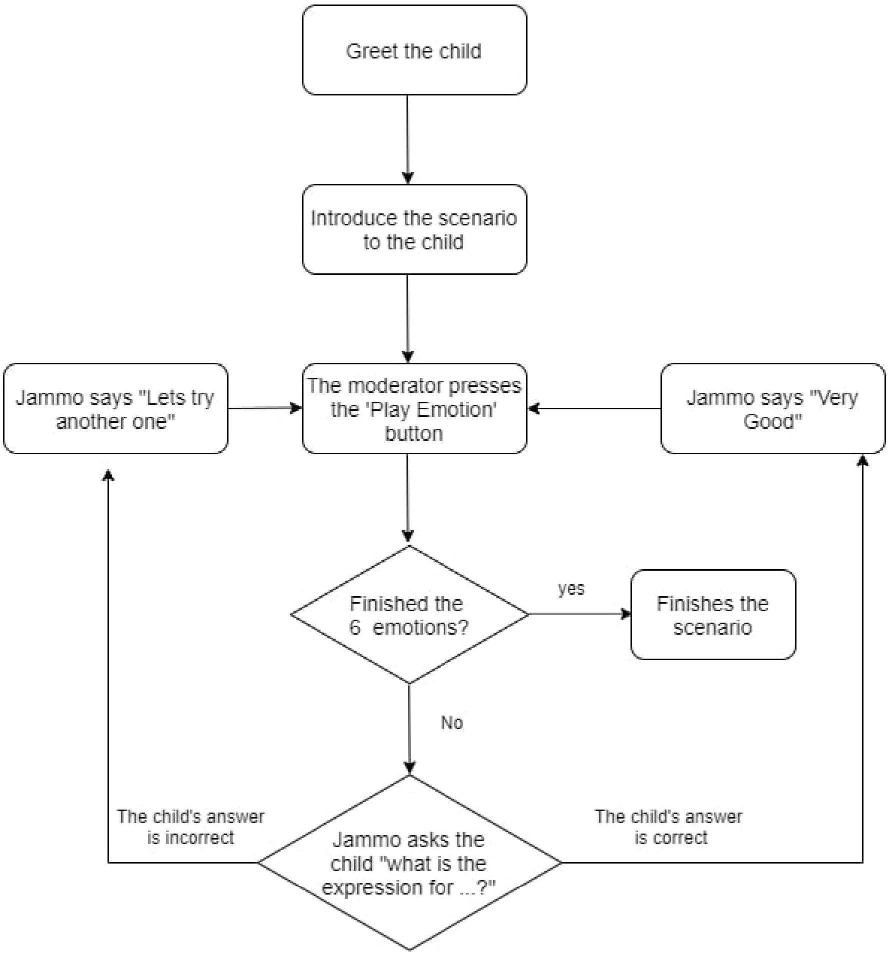


Figure 42: Flowchart of the emotion recognition: Phase II- pretest.

Hello! My name is Jammo Robot.

[Performing "Hello" gesture]

Now! We are going to play an emotion expression game. I am going to tell the name of the emotion, and you are going to express this emotion. The name of the emotion will appear here. Let's start.

[Performing "Talking" gesture]

Then

[Performing "pointing" gesture to the bubble]

The moderator presses " Play Emotion".

What is the expression for "happy"?

[Wait for the child's answer]

If the child's answer is correct !

Very Good!

[Performing "Very Good" gesture]

If the child's answer is incorrect !

Let's try another one!

[Performing "No" gesture]

The moderator presses the "Play Emotion"

After finishing all the six emotions

Now! We have finished. Bye Bye

[Performing "Goodbye" gesture]

Figure 43: Emotion recognition: phase II - pre-test script.

Once the pre-test is done, the child proceeds to the training session, where the Jammo VRobot expresses the six emotions one at a time (e.g. sways from side to side with excitement moving its arms) and says the name of the emotion (e.g. happy). For each emotion, the child is asked to imitate Jammo (e.g. "Now, it is your turn. What is the expression for happy?"). A cartoon character representing this emotion also appears on the screen as shown in Figure 44.

Figure 44: Emotion recognition: Phase II-training.

After completing the four training sessions, the child proceeds to the post-test, which is identical to the pre-test, then a follow-up test after two weeks. The scores in the pre, post, and follow-up tests will be compared.

5.3.2.2 Phase III

Phase III is for training emotion recognition and expression through storytelling. The goal here is to identify the affective state of the Jammo VRobot at the end of a social context. In the pre-test, the Jammo VRobot asks the child to recognise emotion conveyed in the social story. He then narrates a story where he is the main character then asks the child to recognise the emotion being conveyed by saying "How should I feel?". The child is expected to choose one of the three options that appear on separate buttons on the screen as shown in Figure 45. Six stories were designed to represent the six basic emotions as shown in Table 8. Figure 46 shows the flowchart of the scenario.

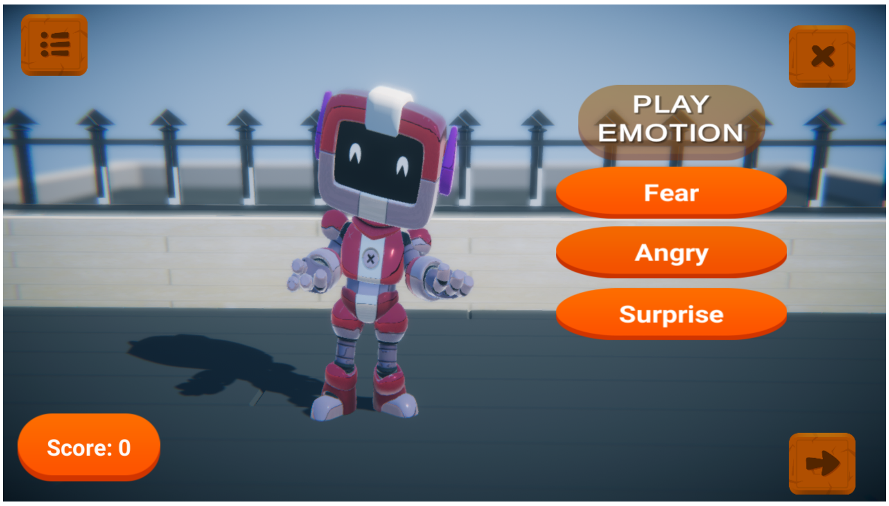


Figure 45: Emotion recognition: Phase III - pretest.

Table 8: Social stories narrated by Jammo VRobot.

Emotion	Social story
Happy	My birthday was yesterday. My parents gave me a present.
Sad	When I was playing. I felt on the floor.
Angry	My friend took my toy without asking me.
Surprise	I came home, and I saw something new. My mum's hair was orange.
Fear	When I was sleeping. I heard something moving in my closet.
Disgust	when I was walking in the park. I stepped on the dog's poop.

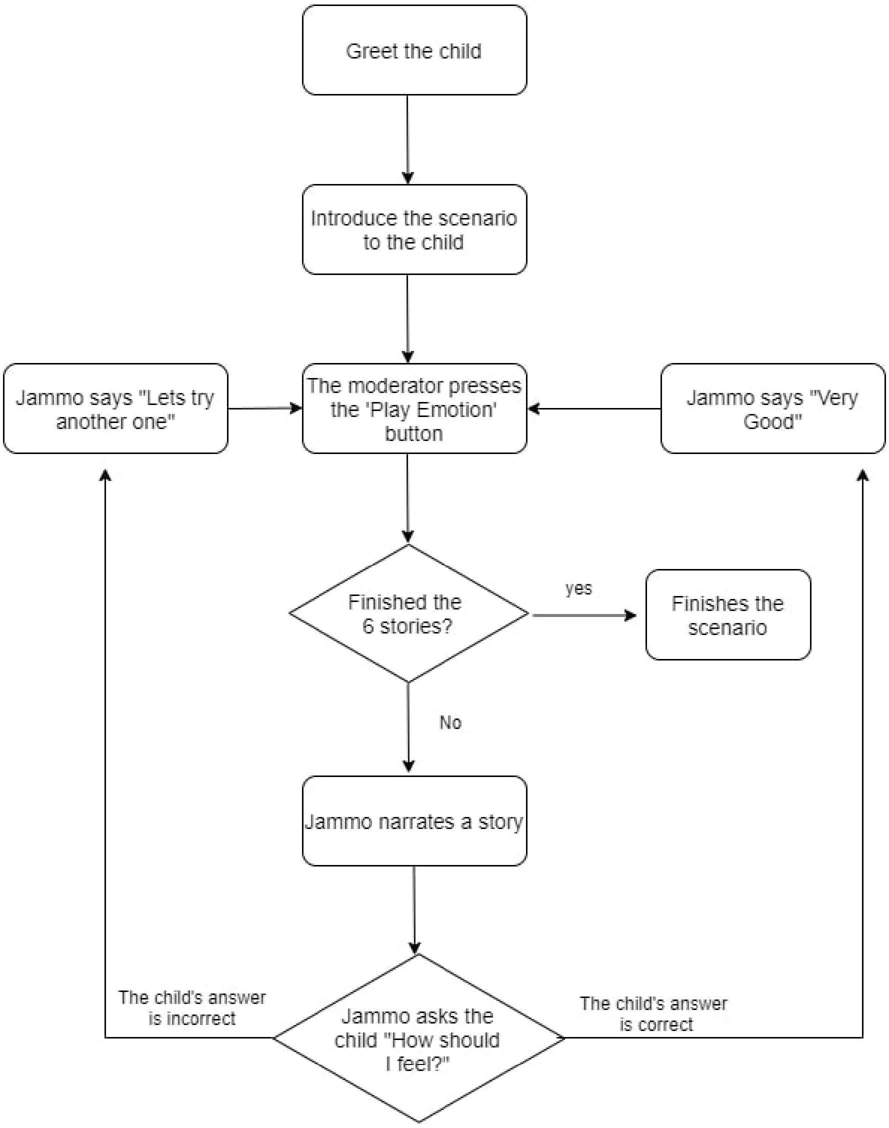


Figure 46: Flowchart of the emotion recognition: Phase III- pretest.

In training, the Jammo VRobot narrates the same stories as the pre-test while expressing the emotion in context. For each story, the child is asked to imitate Jammo's emotions. According to the therapists, children with ASD use vision as their primary input for information (Costa et al., 2014). Thus, the Jammo-VRobot tool incorporates different 3D environments, taking into account the plot of each story. These environments include a playground, birthday party, park, living room, and a bedroom as shown in Figure 47.

Figure 47: The designed social stories.

The post-test and follow-up tests are the same as the pre-test, except the post-tests involve different social stories. Table 9 lists the stories used in the post-tests. As with the previous tests, the scores will be compared.

Table 9: Emotion recognition post-tests social stories.

Emotion	Social story
Happy	Today, my mum took me to the beach. We played on the sand, and we ate ice cream.
Sad	My dad has to go on a trip. How I feel when my dad goes on a trip.
Angry	When we were playing. My sister broke my toy.
Surprise	I thought my friends and my family forgot about my birthday when I got home, they had a party for me.
Fear	When I was watching a TV, I saw a spider on the wall.
Disgust	I took a big bite of an apple. It tasted awful.

5.3.3 Intransitive Gestures Scenarios

Gestures are a prerequisite for the development of languages. They are also an early indication of communication skills. The proposed scenario's structure was adapted and modified from So et al. (2018b,a). The intransitive gesture training programme aims to teach children with HFA to recognise gestures, imitate them, and reproduce them in a social context. The training programme targets 11 gestures used daily and includes: where, stop, not allowed, me, come, awesome, good job, hungry, look at this, yes, and hello. The programme takes the same procedure as the emotion recognition training programme, and it consists of three phases. Each phase includes a pre-test, four training sessions, and two post-tests. The child receives two post-tests (post-test and follow-up post-test) after finishing each phase, one immediately after finishing the training and the other one after two weeks.

5.3.3.1 Phase I

The pre-test scenario aims to evaluate the child's ability to recognise gestures produced by the Jammo VRobot. The scenario starts with the Jammo VRobot greets the child and giving instructions about the pre-test. Jammo VRobot was programmed to perform the 11 gestures randomly. The scenario scene consists of the Jammo VRobot, buttons and score as shown in Figure 48.

Figure 48: Intransitive gesture: pretest-Phase I.

Once Jammo has finished giving the instructions, the moderator/observer clicks the 'Play Gesture' button for him to display a random gesture (e.g. both arms wide open and face upward). He then asks the child to indicate the name of the gesture by saying "What is the meaning of this gesture?". On the scene, three choices are displayed in separate buttons ('Look at this!', 'Me', and 'Where'). The child responds either by clicking the correct button or by saying the meaning of the gesture verbally. Depending on the response, Jammo either praises the child by saying "very good" and performs motivational animation or says "let's try again" and performs 'No' animation.

The moderator/observer then clicks the 'Play Gesture' button to instruct the robot to perform the next gesture. The pre-test is complete when the 11 gestures are covered. The score counts the child's correct answer. The flowchart and the scenario's script are shown in Figures 49 and 50.

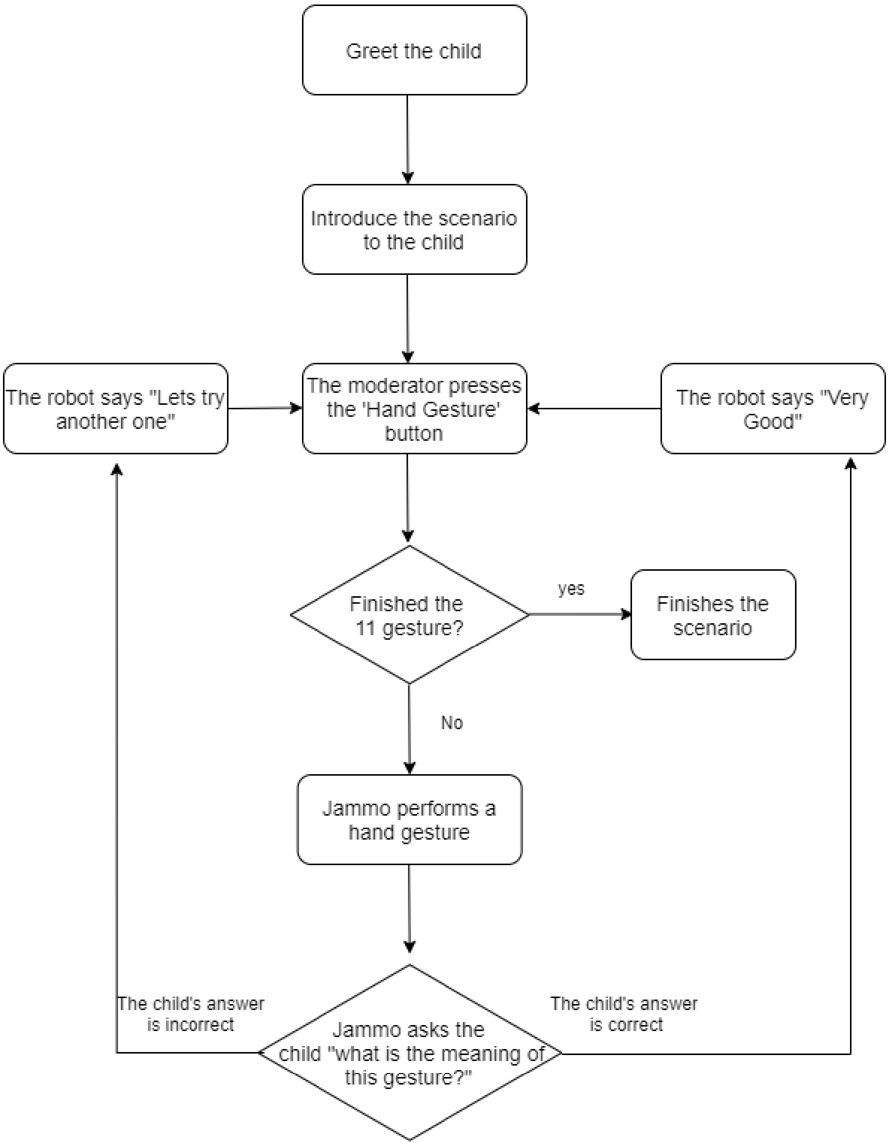


Figure 49: Flowchart of the intransitive gesture: Phase I - pretest.

Hello! My name is Jammo Robot.

[Performing "Hello" gesture]

Now! We are going to play a gesture recognition game. I am going to perform a gesture and you are going to say the name of this gesture. Let's go.

[Performing "Talking" gesture]

The moderator presses " Hand Gesture" to play one of the 11 gestures.

What is the meaning of this gesture?

[displaying a hand gesture]

[Wait for the child's answer]

If the child's answer is correct !

Very Good!

[Performing "Very Good" gesture]

If the child's answer is incorrect !

Let's try another one!

[Performing "No" gesture]

The moderator presses the "Hand Gesture"

After finishing all the 11 gestures.

Now! We have finished. Bye Bye

[Performing "Goodbye" gesture]

Figure 50: Intransitive gesture: phase I - pre-test script.

After finishing the pre-test, the moderator presses the 'Next' button to start the training scenario as shown in Figure 51. There are four training sessions, with two sessions per week. In the training scenario, the Jammo VRobot demonstrates the 11 gestures one at a time randomly (e.g. point at something with the index finger) while indicating the name of that gesture ("Look at this!"). The name of the gesture is displayed in a bubble. The child completes the training after covering the 11 gestures.

Figure 51: Intransitive gesture: Phase I-training.

After all four training sessions are complete, the child progresses to the post-test, which is identical to the pre-test. The follow-up test is taken after two weeks of completing the training sessions.

The scores of the child in the three tests will be compared.

5.3.3.1.1 Phase II

The aim of phase II is to teach the child to produce gestures through imitation. In the pre-test scenario, the Jammo VRobot asks the child to demonstrate a gesture by saying (e.g. "What is the gesture for Where?"). The name of the gesture (e.g. "Where") appears in a bubble as shown in Figure 52. The moderator/observer evaluates the child's gesture and either clicks the 'Correct' or 'Incorrect' button. There is a score for counting the child's correct answers. The scenario's script and the flowchart shown in Figures 53 and 54.

Figure 52: Intransitive gesture: Phase II - pretest.

Hello! My name is Jammo Robot.
[Performing "Hello" gesture]
Now! We are going to play a gesture expression game. I am going to say the name of the gesture and you are going to perform that gesture. The name of the gesture will appear here! Let's go.
[Performing "Talking" gesture]

The moderator presses "Gesture"
What is the gesture for "Where"?

[Wait for the child's answer]
Then
[Points to the bubble]

If the child's answer is correct !
Very Good!
[Performing "Very Good" gesture]

If the child's answer is incorrect !
Let's try another one!
[Performing "No" gesture]
The moderator presses the "Gesture"

After finishing all the 11 gestures.
Now! We have finished. Bye Bye
[Performing "Goodbye" gesture]

Figure 53: Intransitive gesture: phase II - pre-test script.

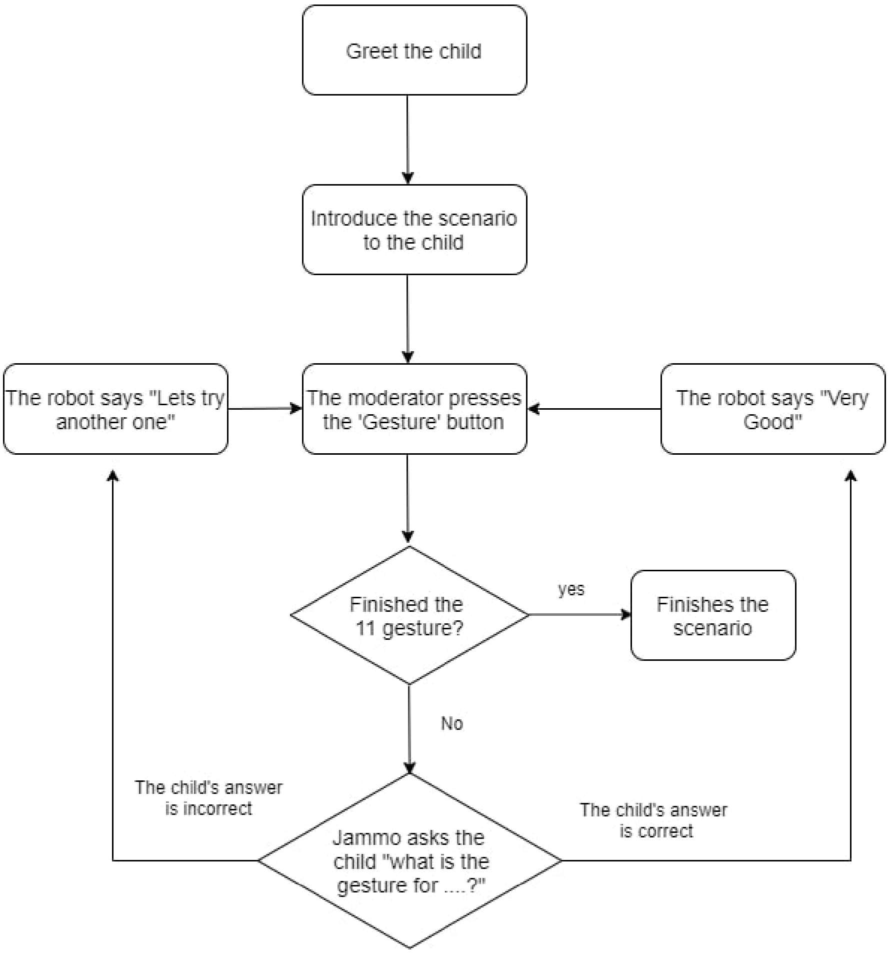


Figure 54: Flowchart of the intransitive gesture: Phase II - pretest.

In the training sessions, the Jammo VRobot animates a gesture (e.g. Point at something with index finger) and says the name of the gesture (e.g. "Look at this"). For each gesture, the child is asked to imitate Jammo (e.g. "Now, it is your turn. What is the gesture for look at this?"). The name of the gesture appears in a bubble as shown in Figure 55. The training session finishes after the 11 gestures were covered.

Figure 55: Intransitive gesture: Phase II - training.

After completing the four training sessions, the child proceeds to the post-test, which is identical to the pre-test, then a follow-up test after two weeks. The scores of the three tests will be compared.

5.3.3.1.2 Phase III

This phase is for teaching gesture recognition and expression through storytelling. The Jammo VRobot asks the child to show the relevant gesture in different social situations during the pre-test session. Jammo VRobot narrates 11 social stories in a randomised order, then asks the child to demonstrate a relevant gesture appropriate gesture by saying "what gesture should I produce?". The moderator/observer evaluates the child's gesture, and either clicks the 'Correct' or 'Incorrect' button. Table 10 lists the stories used in the pre-test.

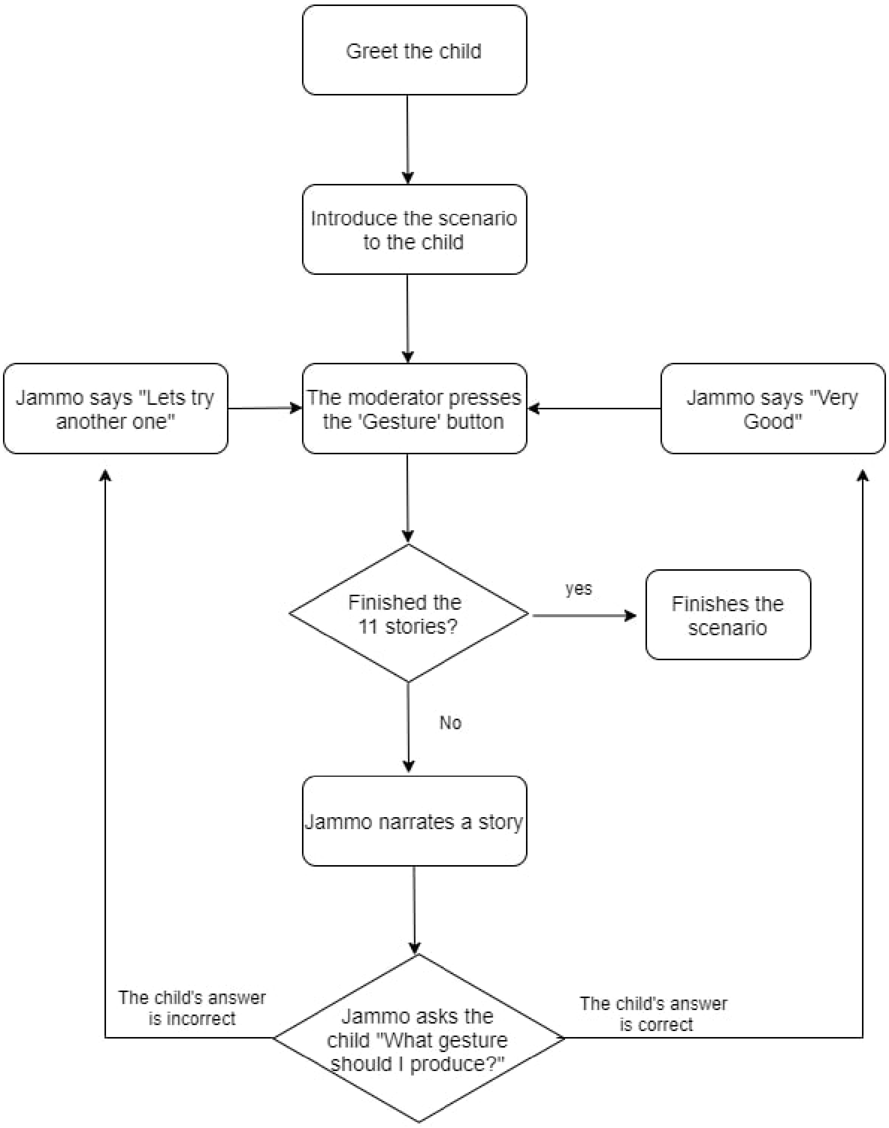


Figure 56: Flowchart of the intransitive gesture: Phase III - pretest.

Table 10: Intransitive gesture pre-test social stories.

Gesture	Social story
Where	I asked my friend where we should go for lunch.
Come	I wanted to show my mum a present. So I asked her to come.
Look at this	I saw a flower. So I said to my dad, look at this.
Hungry	I did not eat my breakfast, so I feel hungry.
Me	Me and my friend go to the park.
Awesome	I won a competition, so I said awesome.
Hello	When I see my friends, I say hello.
Yes	My mum asked me. Have you finished your homework? I said yes.
Good job	My friend did well in the test. So I said good job.
Stop	My friend was crossing the road when the red light is on. So I said stop.
Not allowed	I am not allowed to speak loudly in the library.

The child then progresses to the training sessions, where the Jammo VRobot narrates the same stories told during the pre-test while demonstrating the relevant gesture for each story. After each story, the Jammo VRobot asks the child to imitate the gesture by saying "Now it is your turn". The social stories include a playground, living room, classroom, street, and a library as shown in Figure 57.

Figure 57: Intransitive gesture: social stories.

After completing the 4 training sessions, the child then takes the post-test. Two weeks after the training, the child is expected to take the follow-up test. The social stories in the post-test and follow-up test are different from those told during the pre-test. Table 11 lists the stories used in the post and follow-up tests.

Table 11: Intransitive gesture post-tests social stories.

Gesture	Social story
Where	I was looking for my back bag when I could not find it. I asked my mum where is my back bag.
Come	When I was watching the television, my mum's favourite show started. So I asked her to come.
Look at this	I wanted to show my brother a butterfly, so I told him, look at this.
Hungry	After I finished my sports training. I felt hungry.
Me	Me and my friend wanted to play a video game, but we had to finish the homework first.
Awesome	I have finished my homework, so I felt awesome.
Hello	I was walking in the supermarket with my mum, and I saw my uncle. So I said hello.
Yes	My school friend asked if he could borrow my pencil, so I said yes.
Good job	I was throwing a ball to my dog when he caught it. I said good job.
Stop	When I hear the teacher says the exam is finished. I need to stop writing and close my paper.
Not allowed	I am not allowed to talk in the exam room.

5.4 Conclusion

This chapter presented the Jammo-VRobot environment that deploys the Jammo VRobot to interact with children with HFA. The training utilises a training programme that targets three social skills: imitation, emotion recognition and expression, and intransitive gestures. The training scenario for each skill was adapted from previous work done using physical robots (Huijnen et al., 2017; Costa et al., 2014; So et al., 2018b; Soares et al., 2019; Lux AI, 2019). The structure of the imitation skill training programme replicates the work by Huijnen et al. (2017) and Costa et al. (2014). In the training scenario, Jammo VRobot demonstrates 9 hand gestures and asks the child to imitate them. The emotion recognition and expression training programme built on the work done by (Lux AI, 2019; Soares et al., 2019; Costa et al., 2014). The training programme is divided into three phases, and each phase contains a pre-test, four training sessions, a post-test, and a follow-up post-test. Phase I aims at recognising the basic six emotions and teaching the child when each emotion is most likely felt and the colour representing each emotion. Phase II aims at expressing the basic six emotions through imitation. Phase III is for training emotion recognition and expression through storytelling. The intransitive gesture training programme replicates the work using the NAO robot by (So et al., 2018b). This training programme also consists of three phases. Phase I aims to teach the child to recognise 11 intransitive gestures, phase II aims at expressing these gestures through imitation, and phase III is based on storytelling technique where the child learns to recognise and express gestures in social situations.

The proposed social skill training programme combines more than one training programme in a unified social skill training programme. Thus, parents, teachers, and practitioners can use the proposed training programme to train and enhance several skills of their children with HFA. Jammo-VRobot environment overcomes the limitation of social robots and immersive VE in terms of cost and availability. In case of absence or limited availability of traditional social skills training interventions, social robots, and immersive VE interventions, the Jammo-VRobot environment provides an excellent opportunity for parents to train the social skills of their children with HFA at home. Parents can use the Jammo-VRobot environment anywhere with a computer without needing technicians, professionals, or even without involvement from researchers. Additionally, the Jammo-VRobot environment ensures the repetition of tasks, which is the key to success in training children with HFA. The Jammo-VRobot environment promotes the triadic interaction between the child, parent or teacher, and Jammo VRobot. That type of interaction ensures the involvement of parents and teachers in the training process, which results in improving the outcomes of social skill training interventions. Providing an Arabic version of the Jammo-VRobot environment will increase its availability in countries where English is not the first language. Especially in Egypt, Jammo-VRobot is a promising solution for parents of children with ASD who suffer from the absence of assistive technology interventions and the limited availability of traditional SST interventions. According to Gobrial (2018), most children with ASD in Egypt are given little to no education because they either drop out of mainstream schools or families can not afford the expenses of private schools and specialised centres.

6 Experimental Sessions

The Jammo-VRobot social skill training environment aims to provide an accessible training environment and potentially generalise the taught skills for imitation, emotion recognition and expression, and intransitive gesture skills. An online version of the Jammo-VRobot environment was developed as well as the desktop version.

The evaluation process was conducted mostly online with 11 participants as well as on-site with 4 participants. The participants received 24 sessions, with 2 sessions per week.

6.1 Chapter Overview

The purpose of this study was to evaluate the effectiveness of the Jammo-VRobot environment in training and enhancing three social skills of children with HFA. Additionally, the study aimed to compare the results achieved through the interaction with Jammo VRobot and those achieved through interaction with the physical robots used by the state-of-the-art. Finally, it aimed to assess the potential of generalising the taught skills. As the study aimed to evaluate the effectiveness of the Jammo-VRobot environment, the one-group pre-test and post-test design was employed.

In this chapter, the characteristics of the participants are presented, followed by explaining the experimental setup both online and on-site. Afterwards, the length of the intervention and the procedure of the intervention are discussed. Data collection methods are explained, considering their aim for assessing the participants' social skills before and after the intervention. The evaluation process is divided into three phases: pre-intervention, training, and post-intervention. A discussion of the data collection instruments is presented. Finally, the results obtained from the training sessions and the questionnaires are analysed and reported.

6.2 Method

6.2.1 Participants

To evaluate the Jammo-VRobot training environment, recruiting participants is essential. However, recruiting young children with ASD is challenging. 15 children with HFA, aged between 4-12 years old (M=6.6, SD=2.87), participated in this study, including 10 boys and 5 girls. 10 children aged between 4-8 years, and 5 children aged between 9-12 years.

The inclusion criteria for the participants were as follows: (a) All the children diagnosed as HFA by a specialist and with an IQ >70. (b) Ages between 4-12 years old. Specialists highlight the importance of early intervention for children with ASD, especially regarding their social communication skills and development. Therefore, this study targets young participants. (c) The teacher and parents reported difficulties in imitation, emotion recognition and expression, and intransitive gestures skills. (d) All participants should be verbal and could read simple words to be able to interact with the tool.

The exclusion criteria were as follows: (a) Visual or hearing impairments. (b) Presence of any motor (physical) impairments that may prevent their interaction with the tool.

This research focused only on autism and did not look at any coexisting disorders in the participants. This point could be addressed in the future work.

Ethical approval for the study was obtained from the University of Greenwich Research Ethics Committee (UREC), Appendix A. As part of the study was conducted at a specialised centre in Egypt, ethical approval was also obtained from the Egyptian ministry of education, Appendix B.

It is essential to obtain the parent's consent for interventions that involve young participants. Before getting the parent's consent, they were provided with the information sheet, Appendix H, which outlines the study's purpose and what will be involved during the sessions. Parents were informed that the Jammo-VRobot environment is not suitable for children with visual, hearing, and physical impairments. The parents had the opportunity to ask questions and discuss the study. Concerning the data collection process, data collection tools and data collection process were first explained to the parents. They then were asked to sign the consent for taking part in the study, Appendix I. Parents also had the right to withdraw themselves and their children at any time during the study. All information collected through this study has remained confidential.

Arrangements were made with specialised schools and centres in Egypt and the UK for conducting the live evaluation. A visit to the University of Lincoln's Autism Research Innovation Centre (ARIC) was undertaken in November 2019 to discuss the possibility of recruiting children with ASD involved before in the studies by the university. They organise an annual event (LORIC) for teachers and parents of children with ASD, held on September 2020, where the Jammo-VRobot environment can be presented to the attendees and potential users recruited. The event gathers anyone with an academic, professional, or personal interest in ASD. Parents and teachers of people with ASD come to share their experiences and ideas. It was intended to participate in this event, but the event was cancelled due to the pandemic circumstances (Covid-19) present when this study was conducted. In Egypt, the Egyptian Advance Society for Persons with ASD and Other Disabilities was contacted to plan for the evaluation phase of this research with teachers and parents there with their pupils with ASD. The arrangements and the contacts that were made became no longer available, as the schools and the specialised centres were not accepting any external visitors due to the second wave of Covid-19 in Egypt.

To ensure that the participants with HFA could recognise the gestures produced by the Jammo VRobot in the intransitive gestures training scenarios (discussed in Chapter 5). Typically developing participants (without ASD) were asked to identify the gestures' names demonstrated by Jammo

VRobot in the gesture recognition scenario (phase I, pre-test). The gestures are hello, come, where, look at this, not allowed, yes, hungry, good job, me, stop, and awesome. As described earlier in chapter 5, in the gesture recognition task, Jammo demonstrates 11 gestures randomly one at a time then asks the child to choose one of the three options displayed on separate buttons.

8 Egyptian typically developing children (without ASD) aged 6 to 12 years participated in the preliminary evaluation, including 2 boys and 6 girls. The participants' first language is Arabic, and English is their second language. An observation sheet, Appendix D, was designed for this task and sent to the parents. Parents were asked to observe the child's answers while interacting with the Jammo VRobot then complete the observation sheet and send it back. After receiving the observation sheets, the percentage (%) of the correct answers was calculated as shown in Table 12. The results show that all the 11 gestures had consistency rates of 75% or above, which indicate that the gestures produced by Jammo VRobot are recognisable.

Table 12: List of the 11 gestures producing by Jammo VRobot and their consistency rate.

Form of gesture	**Meaning**	**Consistency rate**
Right/Left palm waving.	Hello	100 %
Head nods up and down.	Yes	100 %
Both arms wide open and face upward.	Where	75%
Hold thumb finger up at the chest level.	Good Job	100%
Right and Left hands hold at the chest level with the palm facing outward.	Stop	100%
Index finger points at something.	Look at this	75%
Index finger points to the chest.	Me	100%
Clapping both hands.	Awesome.	100%
Right/Left arm extend and right/left palm crook and move towards himself.	Come	87%
Right/Left hand move up and down at the lower chest level.	Hungry	87%
Head nods right and left and index finger waving.	Not allowed	100%

6.2.2 Experimental Set-Up

Due to the circumstances of the Covid-19 pandemic and subsequent closure of schools and specialised centres when the study was conducted, the experimentation sessions were conducted mostly online with some on-site. In this subsection, the setup of the environment on-site and online is presented. In both settings, the participants were informed about the training sessions by their parents or teacher to decrease anxiety levels caused by novelty.

6.2.2.1 On-site

The on-site sessions were conducted at a specialised centre (Future Kids Centre) in Egypt. 4 children with HFA (2 boys and 2 girls) participated in the on-site experimental sessions. First, the procedure of the training programme was explained to the teachers besides their role in the experimental sessions. The teacher's role was to administer the sessions by choosing the scenarios, navigating from one scene to another, and assisting the child throughout the session if the child needed assistance. The researcher's role was to observe the sessions and complete the observation sheet (discussed in section 6.2.4.2) regarding the child's activity and behaviour towards the Jammo-VRobot environment. Second, the teacher completed questionnaire (discussed in Section 6.2.4.1) to assess the child's social skills before starting the intervention. Afterwards, the teacher introduced the researcher to the participants, and he then explained what will happen in the sessions. At the completion of the study, the teacher was asked to complete another questionnaire (discussed in Section 6.2.4.3) to evaluate the progress in the child's social skills, the tool's effectiveness, and the tool usability.

Each participant experienced the tool individually, encouraging triadic interaction between the child, the teacher, and Jammo VRobot. The environment setup consists of the child, the teacher, the researcher, and one laptop as shown in Figure 58. The child sits beside the teacher in front of the laptop on a small table. The observer sits at a corner in the room; this allows the researcher to monitor the interaction. Training scenarios were designed to make the training sessions as short as possible to prevent the child from losing concentration.

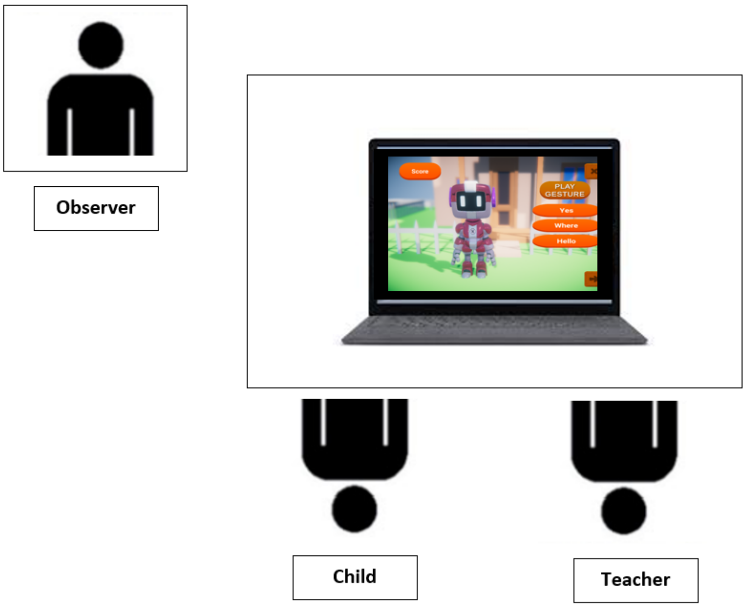


Figure 58: The layout of on-site setup.

The advantage of being on-site is considering the researcher as an observer. The teacher focused on controlling the Jammo-VRobot environment and helped the child and not be overwhelmed by completing the observation sheet and controlling the tool. Additionally, the researcher was objective while filling the observation sheet, unlike the parents that might be subjective towards their children.

6.2.2.2 Online

The online setup took place at the participant's home. An online version of the tool was developed and launched on a website to be available for a wider group and solve the lock-down problem (visit the website here (Abdelmohsen, 2020)). Parents can either access the tool online or download it to their computers. The website contains detailed videos on how to use the tool online and how to download it. Additionally, it explains the training protocol and contains sheets that instruct the parents on what needs to be done for each week. Links to the google sheets that include the questionnaires and the observation sheet were also provided. Figure 59 shows some pages of the website.

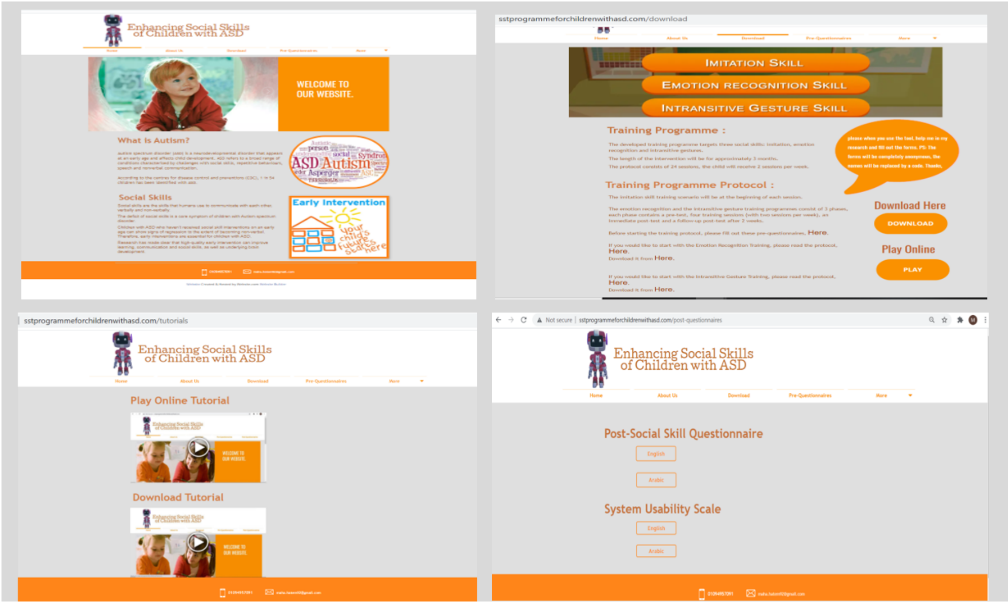


Figure 59: Screenshots of the website.

The website was connected to google analytics to track the users and the new visitors. The website has gone viral, and people from different countries visited it as shown in Figure 60. The website link was posted on several Facebook groups for parents of children with ASD. The nature of the study was explained to parents who decided to participate in the evaluation process through the Zoom application and followed up with them after each session. The parent's role was to control the tool by choosing the scenarios, helping the child, and recording the child's skills and behaviours. Parents only need a laptop (computer) to set up the environment. 11 children with ASD (8 boys and 3 girls) participated with their parents in the online experimental sessions.

Country	Users
	178 % of Total: 100.00% (178)
1. Egypt	71 (39.89%)
2. United States	33 (18.54%)
3. United Kingdom	29 (16.29%)
4. United Arab Emirates	12 (6.74%)
5. Germany	8 (4.49%)
6. Finland	6 (3.37%)
7. France	4 (2.25%)
8. Belgium	3 (1.69%)
9. Austria	2 (1.12%)
10. Greece	2 (1.12%)
11. Netherlands	2 (1.12%)
12. Australia	1 (0.56%)
13. Iraq	1 (0.56%)
14. Japan	1 (0.56%)
15. Montenegro	1 (0.56%)
16. Sweden	1 (0.56%)
17. Vietnam	1 (0.56%)

Figure 60: Google analytics: visitor's country.

6.2.3 Procedure

The intervention programme lasted for approximately three months. The training programme protocol consists of 24 sessions as shown in Figure 61. The participants received two sessions per week. Each session lasted for approximately 15-20 minutes, but sometimes it depended on the child's attention and progress. The first 12 sessions target emotion recognition and expression skills, and the following 12 sessions target intransitive gestures skills. As mentioned earlier, the emotion recognition and expression training programme consists of three phases; each phase contains a pre-test, four training sessions, a post-test, and a follow-up post-test after two weeks. After the first four sessions, the child is expected to recognise the basic six emotions. The following four sessions are expected the child to identify and imitate the basic six emotions. In the last four sessions of the emotion recognition training programme, the child is expected to attribute the relevant emotion to a given context.

Skill	Phase	Week	Session	Focus	Expected outcome
Emotion Recognition	Phase I	1	1-2 session	Emotion recognition	Recognising the basic six emotions.
		2	3-4 session		
	Phase II	3	5-6 session	Emotion imitation	Recognise and imitate emotions.
		4	7-8 session		
	Phase III	5	9-10 session	Context-emotion association	Attribute the proper emotion associated with a given situation.
		6	11-12 session		
Intransitive Gesture	Phase I	7	13-14 session	Gesture recognition	Recognising the 11 gestures.
		8	15-16 session		
	Phase II	9	17-18 session	Gesture imitation	Recognise and imitate 11 gestures.
		10	19-20 session		
	Phase III	11	21-22 session	Context-gesture association	Attribute the proper gesture associated with the given situation.
		12	23-24 session		

Figure 61: Session organisation.

The intransitive gesture training programme followed the same protocol as the emotion recognition and expression training programme. The child is expected to recognise the taught 11 gestures by the end of the first four sessions. The following four sessions are expected the child to recognise and imitate those 11 gestures. In the last four sessions of the training programme, the child is expected to recognise and imitate the gestures in a social context.

The online and on-site settings followed the same protocol.

An Arabic language version of the Jammo-VRobot tool was developed to make it easier for the parents, teachers, and children. Additionally, the Arabic version will increase the availability of the Jammo-VRobot tool in the Middle East where English language is not the first language. Therefore, the questionnaires and the observation sheet were designed in Arabic and English languages. For the Egyptian Arabic language, the used expressions were based on the teachers' consultation at the centre.

6.2.4 Data Collection

The aim of collecting data is to analyse the participant's social skills and evaluate the effectiveness of the Jammo-VRobot environment in improving imitation, emotion recognition and expression, and intransitive gesture skills of the participants. Furthermore, it helps to assess the generalisation and maintenance effects of the environment on the participants. As mentioned in section 3.4, generalisation is the production of behaviour or skill in a context where they were not directly taught, while maintenance is the continuation of performing the taught skill or behaviour even after the training was withdrawn (Gunning et al., 2019). Data can be collected through qualitative and quantitative methods. Qualitative methods include interviews, focus (control) groups, and observations, while quantitative methods include questionnaires, surveys, and records.

In this study, a hybrid method of qualitative and quantitative data collection was used. As a qualitative measure, observation was used, and questionnaires were used as a quantitative measure. The data was secured and only used by the researcher. Codes replaced the participants' names on the questionnaires, and any identifying information was removed.

The evaluation process is divided into three stages: baseline measures (pre-training assessment), intervention measures, and outcomes measures (post-training assessment).

6.2.4.1 Baseline Measures

The baseline measures aim to assess the child's ability in the target skills (imitation, emotion recognition and expression, and intransitive gestures) before starting the intervention (training) sessions to monitor the subsequent change over time when combined with the outcome and the follow-up measures.

As an instrument for data collection, a pre-questionnaire was designed. The Jammo-VRobot social skill training programme targets three social skills: imitation, emotion recognition and expressions, and intransitive gesture. Therefore, the designed questionnaire aims to assess the participants' skills in terms of those three social skills. Some items in the pre-questionnaire were adapted from TRIAD Social Skill Assessment (TSSA) (Wendy Stone et al., 2010). The pre-questionnaire consists of 16 items, Appendix C, 8 items for affective understanding (emotion recognition and expression), 5 items for gesture and play skills, and 3 items to measure the ability of the child in responding to interactions. The affective understanding (emotion recognition and expression) elements evaluate the child's ability to express emotions and understand other's basic emotions. Whereas the gesture and play skills domain looks into the child's ability to understand and use gestures and imitate other's actions. Finally, responding to interaction elements assess the child's ability to concentrate and follow instructions. The questionnaire uses a 4-Point Likert type scale (1 = Never; 2 = Sometimes; 3 = Often; 4 = Almost/Always). The 4-Point Likert scale was used as items from the questionnaire were adapted from TRIAD Social Skill Assessment (TSSA) that used the same Likert scale, and it was sufficient enough to this research, and there was no need to other scales. Additionally, it was used to avoid the safe "Neutral" option because 5-Point Likert scale might not be objective. Furthermore, 5-Point Likert scale might be better for professionals or therapists to diagnose a condition.

Before the intervention sessions start, the teacher and parents were asked to complete social skills pre-questionnaire for their children with HFA.

6.2.4.2 Intervention/Training Measures

The observation was used in the training sessions for collecting data. As mentioned earlier in section 3.4, the screener observes the child's interaction with others, tools, or technology intervention. The live observation was used in this study. For the live observation, an observation sheet, Appendix D, was developed for the screener (parent/researcher) to record the child's skills and behaviours during the training sessions and note the child's response and scores in each test. The designed observation sheet aims to understand the child's behaviour in the sessions and towards the Jammo-VRobot tool. These statements: "Paying attention to the instructions", "Understanding the instructions of the scenario", "Interacting well with the tool", "Responding to the robot questions", and "Repeating words out of context", will give insights about the child's concentration and engagement with the tool. The statements: "Showing positive emotions during the session", and "Showing negative emotions during the sessions" will give insights about the child's emotional demonstration towards the Jammo-VRobot and the tool. The observation sheet also includes some questions regarding the training scenarios, such as "What are the emotions that he/she could not recognise" and "The score". These questions will help the evaluation process in understanding the emotions and gestures that were difficult for the children to recognise or express and compare the scores between the pre-test, post-test, and follow-up test to assess the effectiveness of the Jammo-VRobot environment.

6.2.4.3 Post-training Measures

The post-training measures aim to assess the changes in the child's skills and behaviour, that were measured in the pre-questionnaire, after receiving the training. The effect of the training was assessed through pre and post comparisons. After the completion of the training sessions, the teacher and parents were asked to complete the post-questionnaire. The post-questionnaire contains extra items not included in the pre-questionnaire, Appendix E. These extra items aim to assess the opinions of the teacher and parents about the training programme their children received. For example: "The emotion recognition scenarios used were appropriate to train the child's emotional ability" and "I am satisfied with the emotion recognition scenarios that my child practised". These statements give insight about the parents' and teacher's acceptability of the intervention programme, as their acceptability helps improve and generalise the skills. It also makes them keen to practise the learnt skills with their children at home. The post-questionnaire also contains questions to assess the generalisation of the learnt skills.

To evaluate the parent/teacher satisfaction with the technical aspects of the Jammo-VRobot environment and the tool usability, the System Usability Scale (SUS) (Valencia et al., 2019; Kobak et al., 2011; Miskam et al., 2015) was used. It obtains quantitative feedback on a 0-100 scale regarding the effectiveness, efficiency, and user satisfaction experience while interacting with the tool. The System Usability Scale (SUS) consists of ten items using a five-point Likert type scale, Appendix F, (1 = Strongly Disagree; 2 = Disagree; 3 = Neutral; 4 = Agree; 5 = Strongly Agree) (Valencia et al., 2019). Users ranked each of the 10 questions from 1 to 5 based on their level of agreement. For each odd (Positive) question, subtract 1 from the score, and for the even (negative) questions, subtract their value from 5. In the end, add the total score, then multiply it by 2.5. Table 13 shows the general guidelines on SUS score interpretation.

Table 13: SUS score interpretation.

SUS Score	**Grade**	**Adjective Rating**
>80.3	A	Excellent
68-80.3	B	Good
68	C	Okay
51-68	D	Poor

6.3 Statistical Analysis

Data was analysed using statistical software SPSS (MedCalc, 2021). According to the literature, there are mixed opinions regarding using parametric or non-parametric tests for statistical analysis. Some researchers cautioned against the use of parametric tests when the sample size is small (Marino et al., 2020; Soares et al., 2019), whereas other suggested that using parametric tests is feasible in such a case (Reyes et al., 2019; De Winter, 2013). Paired t-test for a parametric test and Wilcoxon for a non-parametric test are the appropriate methods for comparing the means of dependant data (within the same group; performed by the same child) before and after an intervention to evaluate the significance of improvement. Both tests were conducted to evaluate the significant difference between the pre-test, post-test, and follow-up test scores. The results from both tests were the same among all the phases. Due to the small sample size, the non-parametric (Wilcoxon) test results were reported to be more conservative in the analysis. To determine the statistical significance, the p-value threshold was set to 0.05. Authors refer to statistically significant when the p-value is less than 0.05 ($P<0.05$) (Greenland et al., 2016; So et al., 2018b; Statsdirect, 2021).

6.4 Results

6.4.1 Results From Training Sessions: Intransitive Gesture Training Programme

6.4.1.1 Phase I: Gesture Recognition

This phase aims at examining the learning outcomes of the gesture recognition skills of the participants. As mentioned earlier, the pre-test, post-test, and follow-up test are identical. The Jammo VRobot shows the 11 gestures one at a time in a randomised order while asking the child "What is the meaning of this gesture?". The achievement of this goal was measured by comparing the score of the participants in the pre-test, post-test, and follow-up (delayed) test. In the three tests, the number of questions each child correctly recognised the name of the gesture were counted. The maximum score each participant could achieve in each test is 11, as each test includes 11 questions. Figure 62 shows the number of correct answers each child got in each of the pre-test, post-test, and follow-up test. The comparison between the scores in each test shows that the gesture recognition skill of the participants progressively developed over the sessions. The participants maintained positive learning outcomes two weeks after the training in the follow-up test.

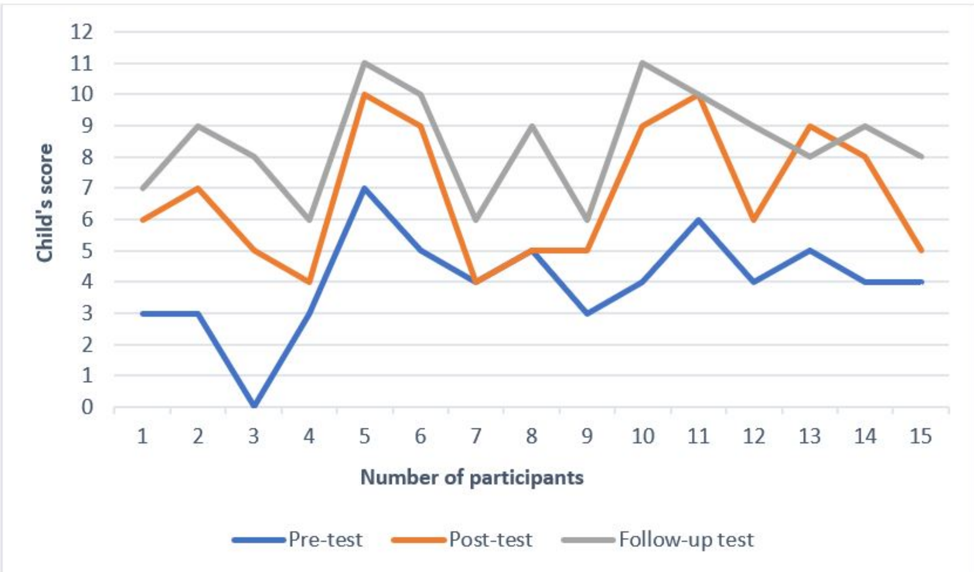


Figure 62: Comparison between the scores of the participants in the pre, post, and follow-up tests: phase I of the intransitive gesture programme.

Afterwards, the mean number of the correct responses across all the participants in each test was calculated as shown in Figure 63. The mean numbers of correct answers in the pre-test, post-test, and follow-up test are 4, 6.8, and 8.4, respectively. The participants' performance in recognising the 11 gestures increased by 25.5% in comparison between the pre-test and the post-test (from 36.3% in the pre-test to 61.8% in the post-test). 40.6% improvement in the participants' recognition skills in comparing the pre-test with the follow-up test (from 36.3% in the pre-test to 76.9% in the follow-up test).

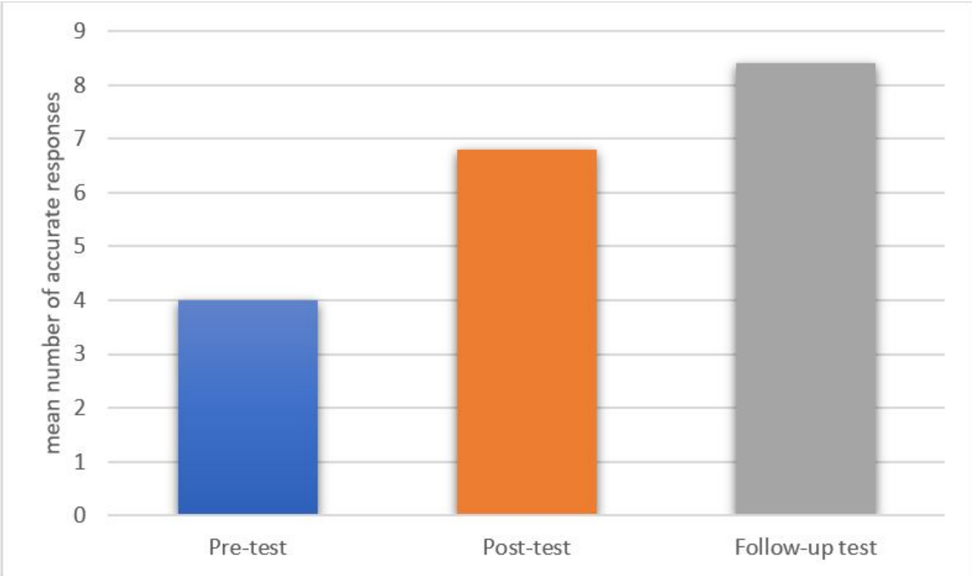


Figure 63: Mean number of accurate responses in the pre, post, and follow-up tests: Phase I of the intransitive gesture programme.

Wilcoxon test was employed for the pre-test and post-test scores to assess the significance of difference after receiving the four training sessions.

- Q1: Is there a significant improvement in the participants' gesture recognition skills after receiving the four training sessions?

  Two hypotheses derive from this research question.

  - H0: There is no significant improvement in the gesture recognition skills of the participants after receiving the training sessions.
  - H1: There is a significant improvement in the gesture recognition skills of the participants after receiving the training sessions.

The results obtained from the Wilcoxon test method are shown in Table 14. The results highlight the p-value between the pre-test and post-test, $P=0.001$ (where $P<0.05$). Therefore, the hypothesis (H0) was rejected, and the hypothesis (H1) was accepted. Thus, it concludes that the intransitive gesture skills of the children significantly improved after receiving the four training sessions. Additionally, the comparison showed that the post-test and follow-up test scores were generally higher than the pre-test scores. The p-value between the pre-test and the follow-up test indicates highly significant differences in the gesture recognition skills of the participants, $p<0.001$.

Table 14: Wilcoxon test results: Intransitive gesture-Phase I.

	Pre-test	**Post-test**	**Follow-up test**
Sample Size	15	15	15
Mean	4	6.8	8.4
Standard Deviation	1.60	2.1	1.68
P-Value (pre-test and post-test)	P=0.001, ($p<0.05$)		
P-Value (post-test and follow-up test)		P=0.002, ($p<0.05$)	

6.4.1.2 Phase II: Gesture Production

Phase II aims to examine the learning outcomes of the gesture production skills of the participants through imitation. In the pre-test, post-test, and follow-up test, Jammo VRobot asks the child to show the 11 gestures one at a time in a randomised order by saying (e.g. "What is the gesture for hello?"). The achievement of this goal was measured by counting the number of times the participants produced gestures correctly in each of the three tests according to four parameters.

- **Hand shape**: open palm vs curled palm vs fist
  - Curled palm when producing HELLO gesture considered incorrect.
- **Direction of movement**: head nods vs head shakes; moving the hand from left to right vs moving it up and down
  - Head shakes while producing YES gesture considered incorrect.
  - Moving hand up and down while producing NOT ALLOWED gesture is considered incorrect.
- **Placement**: hand placed at the chest vs at the head
  - Hand placed at the head while producing ME gesture is considered incorrect.
- **Use of hands**: use one hand vs two hands
  - Use one hand while producing AWESOME gesture is considered incorrect.
  - Use of one hand while producing STOP gesture is considered incorrect.

Figure 64 shows the score of each child in the pre-test, post-test, and follow-up test. The comparison between the scores indicates that the participants were more likely to imitate and produce gestures after the four training sessions. Furthermore, the follow-up test scores suggest that the participants maintained the improvement two weeks after completing the training.

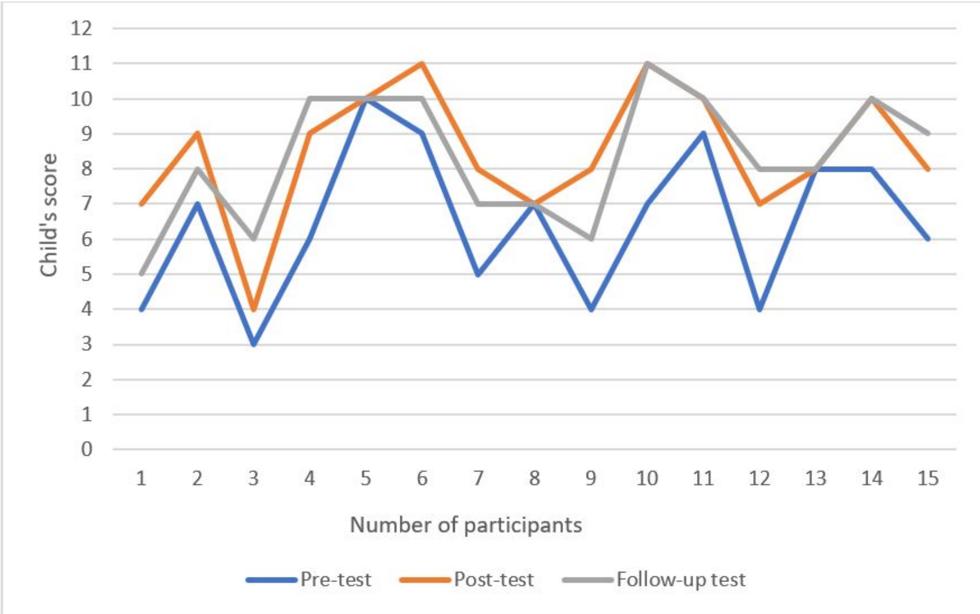


Figure 64: Comparison between the scores of the participants in the pre, post, and follow-up tests: phase II of the intransitive gesture programme.

Subsequently, the mean score of the participants in the three tests was calculated. The mean scores in the pre-test, post-test, and follow-up test are 6.4, 8.4, and 8.3 as shown in Figure 65. In this phase, the children's performance in imitating and producing the taught 11 gestures increased by 18.2% (from 58.70% in the pre-test to 76.90%). The percentage of improvement in the participants' skills in this phase (II) was less than the increase in the gesture recognition phase (I). The participants' performance in phase II - pre-test was already close to 60 %, which gave a small room for progress. The participants' scores in phase II - pre-test were relatively higher than those in phase I - pre-test.

Interestingly, during the four training sessions in phase I, it was observed that the children were imitating the Jammo VRobot and verbally repeating the name of the gestures while teaching them the 11 gestures without asking them to imitate. Therefore, phase II was easier for them than phase I. Additionally, it provides proof of the participants' engagement while interacting with Jammo VRobot.

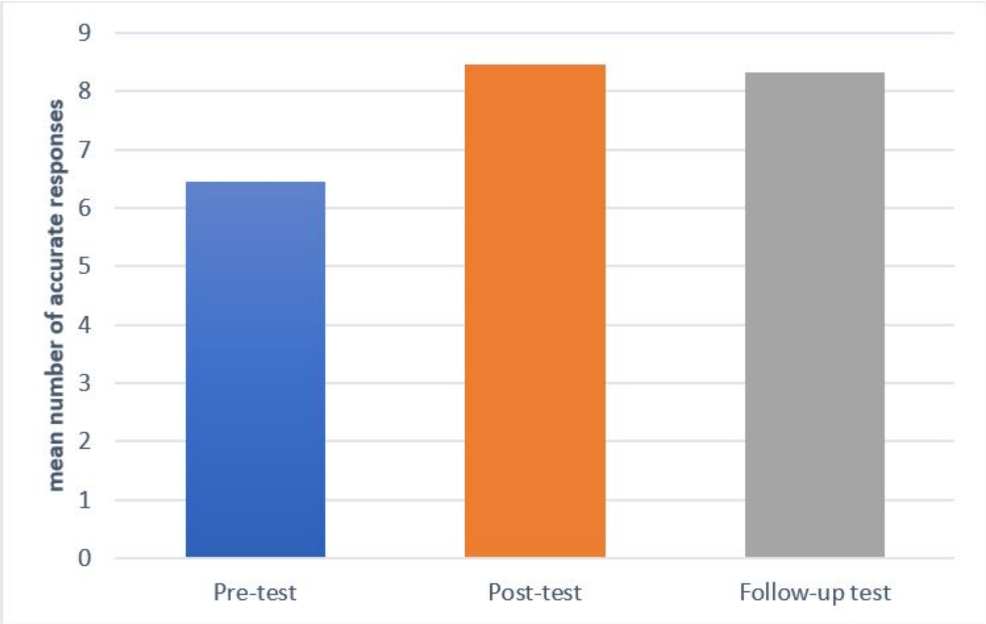


Figure 65: Mean number of accurate responses in the pre, post, and follow-up tests: Phase II of the intransitive gesture programme.

Using a Wilcoxon statistical test as shown in Table 15, the pre-test, post-test, and follow-up test were compared to answer the following research question.

- Q1: Is there a significant improvement in the participants' gesture production skills after receiving the four training sessions?

  Two hypotheses derive from this research question.

  - H0: There is no significant improvement in the gesture production skills of the participants after receiving the training sessions.
  - H1: There is a significant improvement in the gesture production skills of the participants after receiving the training sessions.

The results show a significant difference between the pre-test and post-test, as P=0.002 ($P<0.05$). There was no significant difference between the post-test and the follow-up test (P=0.6). Therefore, the hypothesis (H0) was rejected, and the hypothesis (H1) was accepted. Thus, it concludes that the gesture production skills of the participants improved after receiving the training sessions.

Table 15: Wilcoxon test results: Intransitive gesture-Phase II.

	Pre-test	**Post-test**	**Follow-up test**
Sample Size	15	15	15
Mean	6.4	8.4	8.3
Standard Deviation	2.13	1.84	1.83
P-Value (pre-test and post-test)	P=0.002, (p<0.05)		
P-Value (post-test and follow-up test)		P=0.6, (p>0.05)	

6.4.1.3 Phase III: Gesture Recognition and Production

Phase III aims at recognising and producing the relevant gesture in different social contexts. In the pre-test, Jammo VRobot narrates 11 social stories. He then asks the children to identify and produce the relevant gesture for each story. The post-test and follow-up test involve 11 different stories from those used in the pre-test. The moderator (parent/teacher) judged the accuracy of the gesture production based on the mentioned four parameters in the previous section: hand shape, the direction of movement, placement, and use of hands.

The participants' scores in the pre-test, post-test, and follow-up test show that phase III was the most challenging task for the participants as shown in Figure 66. Nonetheless, after receiving the training sessions, the children got higher scores in the untrained stories (post-test).

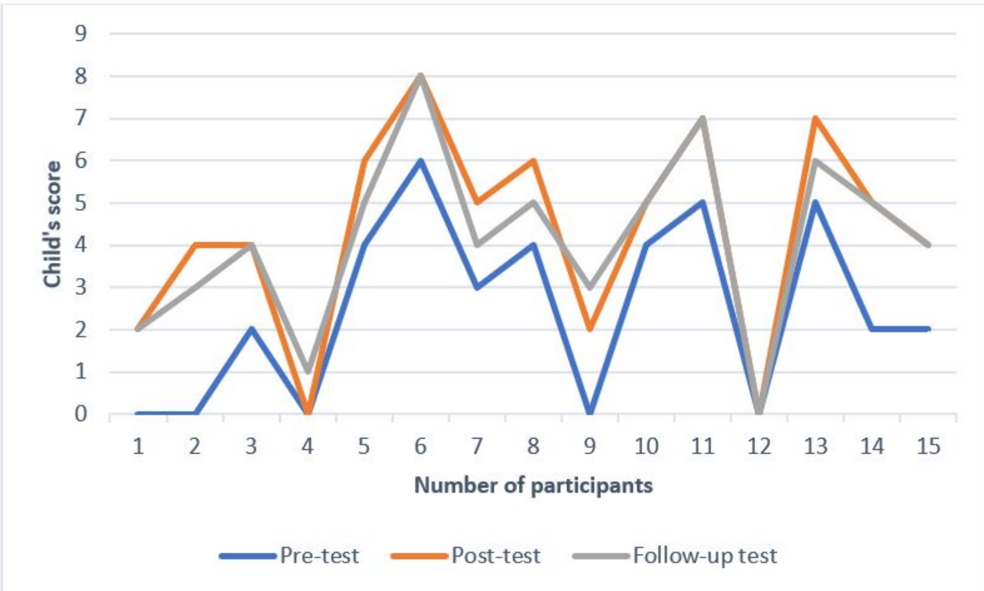


Figure 66: Comparison between the scores of the participants in the pre, post, and follow-up tests: phase III of the intransitive gesture programme.

The mean numbers of the correct answers in each of the three tests are 2.4, 4.3, and 4.1 as shown in Figure 67. However, participants found the storytelling task the hardest; their performance increased by 16.9% in the untrained scenarios (from 22.40% in the pre-test to 39.30% in the post-test). It is noteworthy that the content of the stories in the post-test and follow-up test differs from those used in the pre-test and the training, suggesting that the participants could generalise the gained skills to new social contexts (untrained scenarios). Additionally, the gained skills were maintained two weeks after the training.

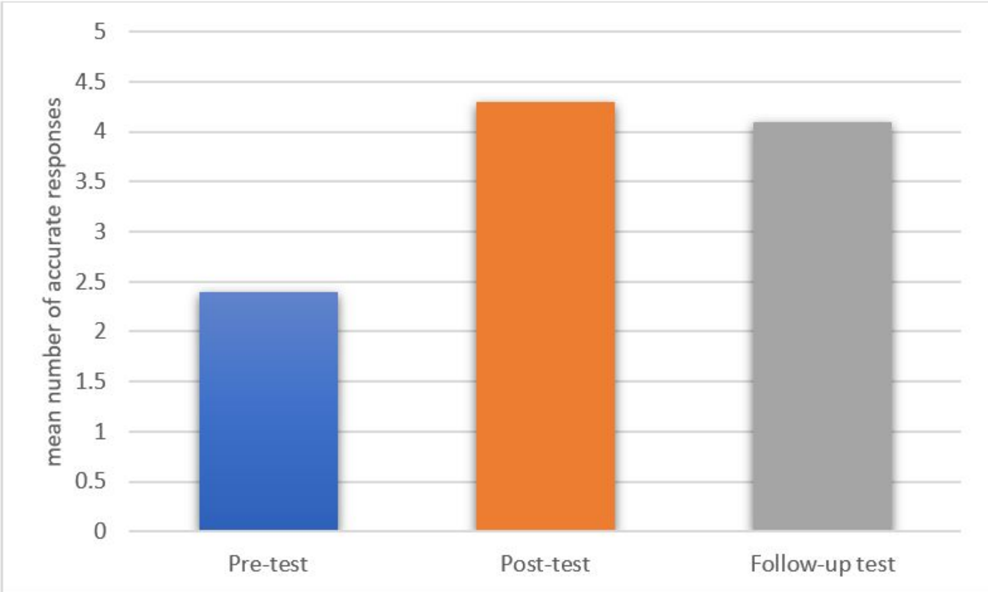


Figure 67: Mean number of accurate responses in the pre, post, and follow-up tests: Phase III of the intransitive gesture programme.

In the analysis of the outcomes using the Wilcoxon statistical test as shown in Table 16, a significant difference was noticed between the pre-test and the post-test where the p-value is less than 0.001. The following research question was answered.

- Q1: Is there a significant improvement in the participants' gesture recognition and production skills in different social contexts after receiving the four training sessions?

  Two hypotheses derive from this research question.

  - H0: There is no significant improvement in the participants' gesture recognition and production skills in untrained social contexts after receiving the training sessions.
  - H1: There is a significant improvement in the participants' gesture recognition and production skills in untrained social contexts after receiving the training sessions.

The results show a significant difference in the participants' skills after receiving the four training sessions. The participants were able to identify and show the relevant gestures in non-training social stories. Thus, hypothesis (H0) was rejected, and hypothesis (H1) was accepted.

Table 16: Wilcoxon test results: Intransitive gesture-Phase III.

	Pre-test	**Post-test**	**Follow-up test**
Sample Size	15	15	15
Mean	2.4	4.3	4.1
Standard Deviation	2.13	2.44	2.13
P-Value (pre-test and post-test)	P<0.001		
P-Value (post-test and follow-up test)		P=0.2, (p>0.05)	

6.4.2 Results From The Training sessions: Emotion Recognition and Expression Training Programme

6.4.2.1 Phase I: Emotion recognition

Phase I provides training for emotion recognition and focuses on the basic six emotions: happiness, anger, sadness, fear, surprise, and disgust. The pre-test, post-test, and follow-up test are identical. The Jammo VRobot expresses the six emotions one after another randomly, then asks the child "What is the name of this emotion?". The maximum score each participant could achieve in each test is 6, as each test involves 6 questions.

Figure 68 shows the score each participant achieved in the pre-test, post-test, and follow-up test. After receiving the four training sessions, the participants' scores suggest progress in the participants' emotion recognition skills. Incorporating cartoon characters with each emotion, besides the gestures, helped to improve the learning process in the training sessions. It kept the participants motivated and paid attention to the information presented, as children with ASD are visual learners.

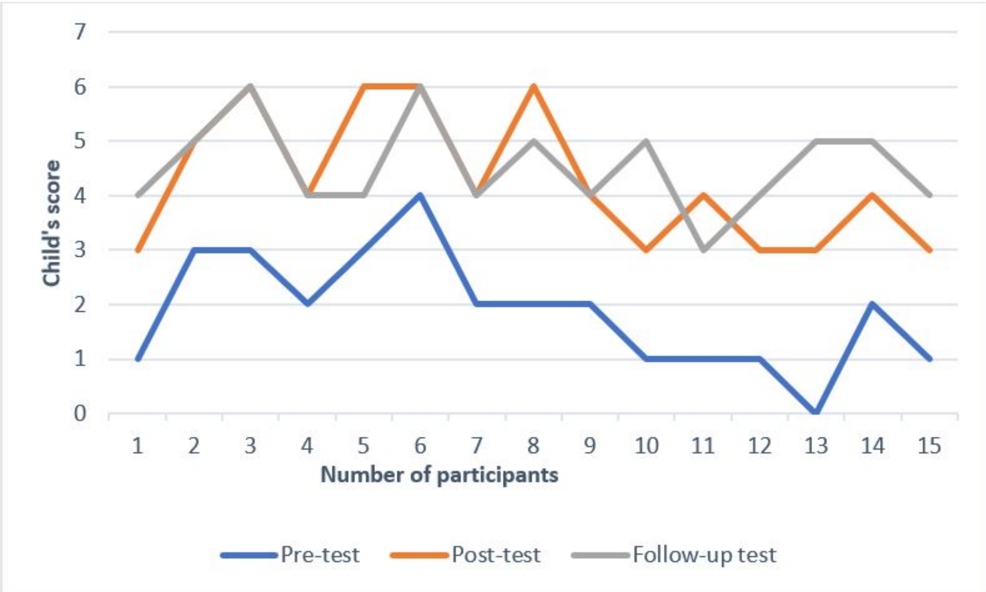


Figure 68: Comparison between the scores of the participants in the pre, post, and followup tests: phase I of the emotion recognition programme.

The mean numbers of correct answers in recognising the basic six emotions in each test were calculated as shown in Figure 69. The results show a 40% increase in the participants' emotion recognition skills (from 31% in the pre-test to 71% in the post-test). The positive learning outcomes were maintained two weeks after receiving the training. The participants' performance in recognising the emotions increased by 44% (from 31% in the pre-test to 75% in the follow-up test).

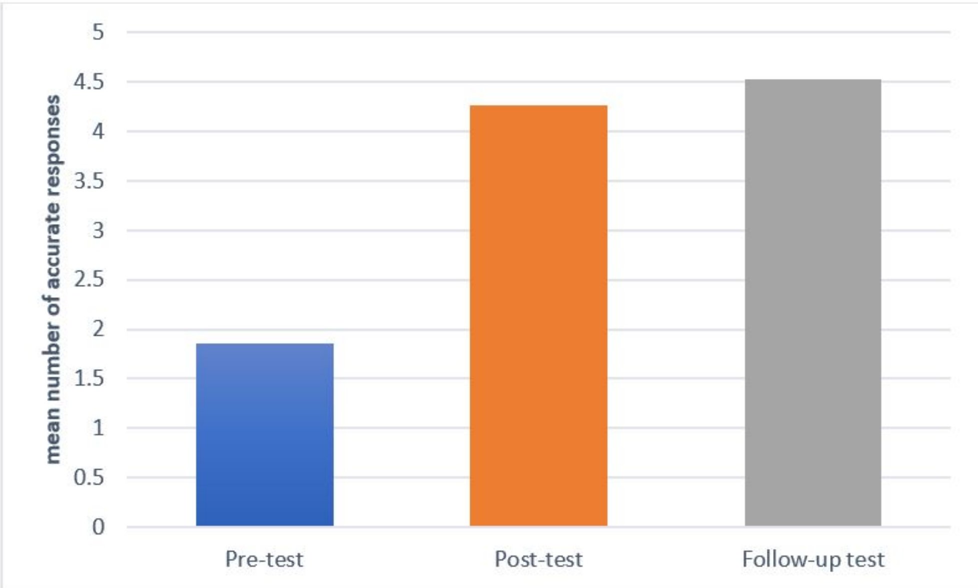


Figure 69: Mean number of accurate responses in the pre, post, and follow-up tests: Phase I of the emotion recognition programme.

Some emotions were easy to recognise than others. Therefore, the proportions of participants providing correct answers to each emotion in this phase were calculated as shown in Table 17. The number of participants who correctly identify HAPPY, SAD, and ANGRY emotion were more significant than those who recognised SURPRISED, DISGUST, and FEAR emotion.

Table 17: The participants' mean score in phase I for each emotion.

Emotion	**Mean score**
Happy	0.86
Sad	0.82
Angry	0.8
Surprise	0.33
Disgust	0.26
Fear	0.51

Figure 70 shows the accuracy percentage of the recognised six emotions expressed by the Jammo VRobot. Only 26.6% and 33.3% of the participants correctly recognised DISGUST and SURPRISE emotion. 86.6% of the participants correctly recognised HAPPY emotion, followed by 82.2% for the SAD emotion. The percentages of recognising ANGRY and FEAR emotions were 80% and 51%, respectively. The confusion in recognising SURPRISE and DISGUST emotions might be due to the importance of facial expressions in showing those emotions. It might be hard to reflect these two emotions without using actuated eyebrows and mouth, which the Jammo VRobot lacks. The result is consistent with some evidence that facial expressions are essential for recognising some emotions (Ghafurian et al., 2020). Furthermore, these results support previous studies that indicate DISGUST and FEAR emotions are difficult for children with ASD to recognise (Askari, 2018). It is also consistent with some evidence that children with ASD have particular difficulty recognising basic negative emotions (Raisingchildren, 2020; Askari, 2018; Rump et al., 2009).

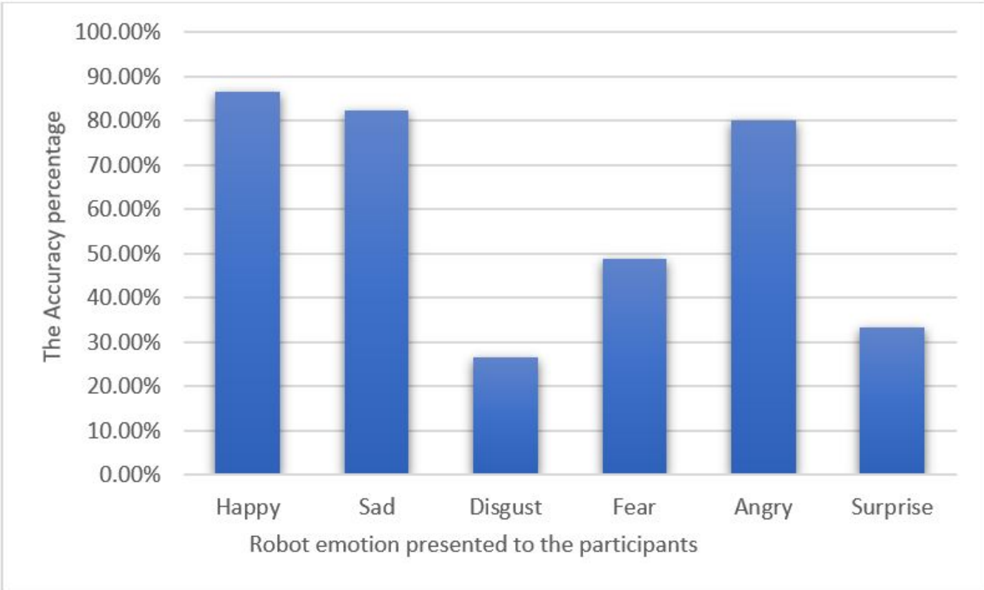


Figure 70: The accuracy percentage of the recognised emotions expressed by the Jammo VRobot in Phase I.

The collected data was statistically analysed using Wilcoxon statistical test as shown in Table 18, to answer the following research question.

- Q1: Is there a significant improvement in the participants' emotion recognition skills after the intervention?

  Two hypotheses derive from this research question.

  - H0: There is no significant improvement in the emotion recognition skills of the participants after receiving the training sessions.
  - H1: There is a significant improvement in the emotion recognition skills of the participants after receiving the training sessions.

Table 18: Wilcoxon test results: Emotion recognition-phase I.

	Pre-test	**Post-test**	**Follow-up test**
Sample Size	15	15	15
Mean	1.87	4.27	4.53
Standard Deviation	1.06	1.22	0.83
P-Value (pre-test and post-test)	P<0.001		
P-Value (post-test and follow-up test)		P=0.3, (p>0.05)	

The results highlight that the p-value was less than 0.001 ($p<0.001$); therefore, hypothesis (H0) was rejected, and hypothesis (H1) was accepted. There was no difference between the post-test and the follow-up test, as the p-value was equal to 0.3 ($p>0.05$). The results conclude that the Jammo VRobot contributes to developing the emotion recognition skills of the participants.

6.4.2.2 Phase II: Emotion expression

This phase aims at exploring the learning outcomes of the emotion expression skill of the participants through imitation. The Jammo VRobot expresses each emotion then asks the child to imitate the expressed emotion. The pre-test, post-test, and follow-up test are identical. The Jammo VRobot asks the participant to express the basic six emotions one at a time in a randomised order by saying "What is the expression for HAPPY?". The moderator judged the accuracy of the expressed emotion.

To test the hypotheses of this phase, the participants' scores in the pre-test, post-test, and follow-up test were recorded. It was observed that all children had better scores in the post-test and the follow-up test than the pre-test as shown in Figure 71. In the pre-test, 46.6% of the participants correctly expressed three of the six emotions, 33.3% expressed four emotions correctly, and 20% expressed two emotions. In the post-test, 13.3% of the children correctly expressed the six emotions, 40% expressed four emotions correctly, and 46.6% of the participants correctly expressed five emotions. After two weeks of the training (follow-up test), the emotion expression skills of the participants were maintained as more than 53% of the participants correctly expressed five emotions. The results emphasise that imitation is a crucial aspect of skill development.

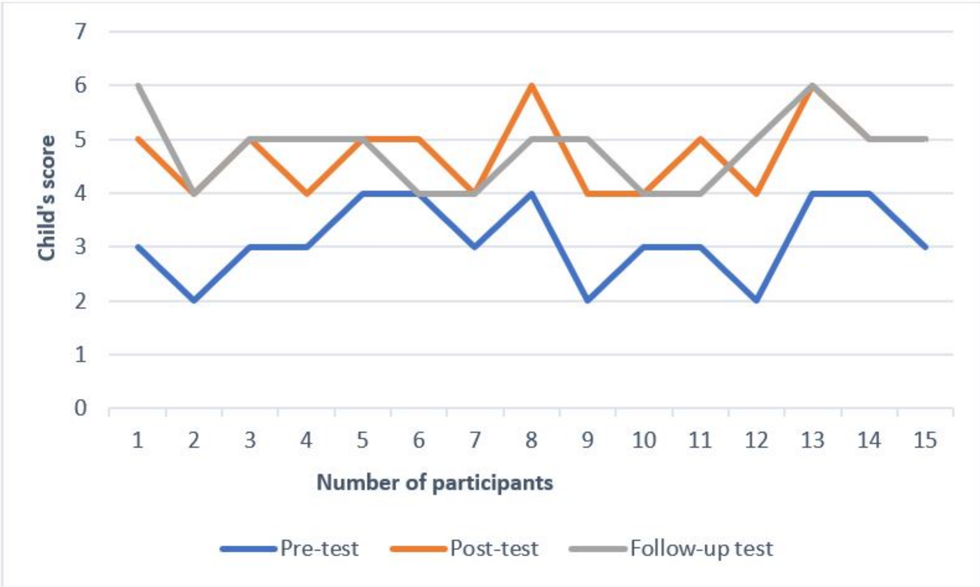


Figure 71: Comparison between the scores of the participants in the pre, post, and followup tests: phase II of the emotion recognition programme.

Figure 72 shows the mean score of the participants in the three tests. On average, the emotion expression skills of the participants increased by 26.8% (from 52% in the pre-test to 78.80% in the post-test), while their skills improved by 28% in the follow-up test (from 52% in the pre-test to 80% in the follow-up test).

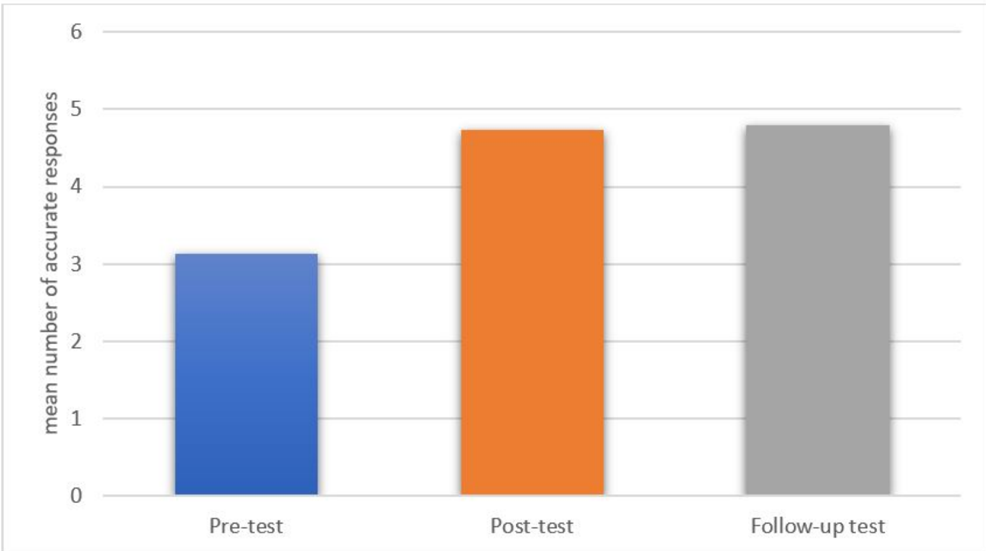


Figure 72: Mean number of accurate responses in the pre, post, and follow-up tests: Phase II of the emotion recognition programme.

Some emotions were difficult to express than others as shown in Table 19. HAPPY and SAD were the most well-expressed emotions, as all the participants expressed them correctly. 80% of the participants correctly expressed the ANGRY emotion, followed by FEAR emotion with 66.6% as shown in Figure 73. The participants found difficulty in expressing SURPRISE and DISGUST emotions.

Table 19: Participants' mean score in phase II for each emotion.

Emotion	Mean score
Happy	1
Sad	1
Angry	0.8
Surprise	0.42
Disgust	0.33
Fear	0.66

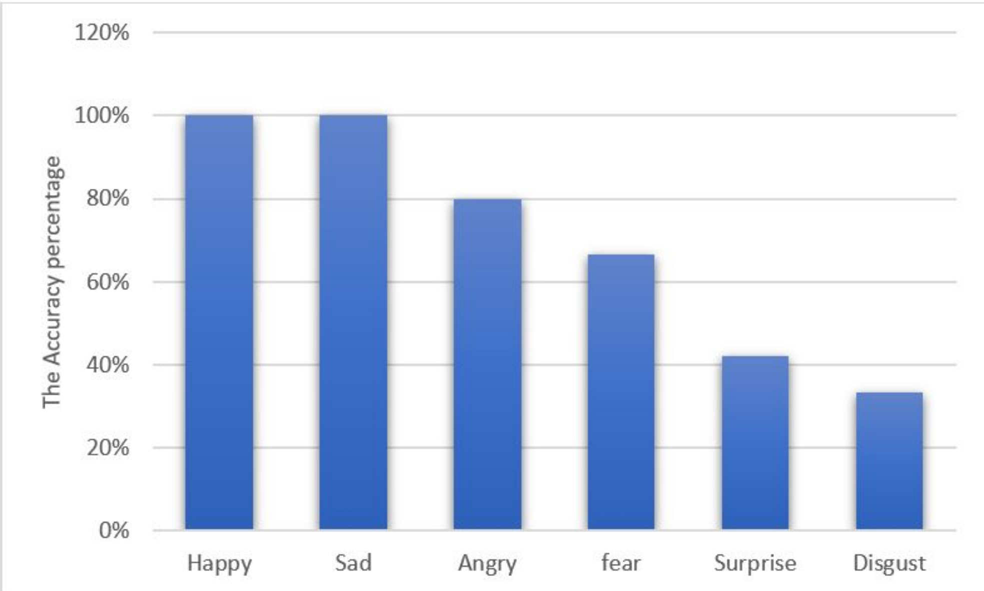


Figure 73: Accuracy percentage of the expressed emotions by the participants in Phase II.

Table 20 illustrates the mean, the corresponding standard deviation (SD) of the correct answers given by the participants in the three tests, and the p-value.

Table 20: Wilcoxon test results: Emotion recognition-phase II.

	Pre-test	Post-test	Follow-up test
Sample Size	15	15	15
Mean	3.13	4.73	4.80
Standard Deviation	0.74	0.70	0.67
P-Value (pre-test and post-test)	P<0.001		
P-Value (post-test and follow-up test)		P=0.7, (p>0.05)	

- Q1: Is there a significant improvement in the participants' emotion expression skills after the intervention?

Two hypotheses derive from this research question.

– H0: There is no significant improvement in the emotion expression skills of the participants after receiving the training sessions.

– H1: There is a significant improvement in the emotion expression skills of the participants after receiving the training sessions.

A significant difference was found when comparing the pre-test and the post-test ($P<0.001$). Therefore, hypothesis (H0) was rejected, and hypothesis (H1) was accepted. The results conclude that the Jammo VRobot contributes to developing the emotion expression skills of the participants through imitation.

6.4.2.3 Phase III: Emotion recognition and expression

This phase evaluates whether the Jammo VRobot can help children with HFA to identify emotions from social situations. Jammo VRobot narrates six emotional stories then asks the children to recognise the relevant emotion at the end of each story. The stories in the post-test and the follow-up test are different from those used in the pre-test. The aim of designing different stories in the post-test is to evaluate the emotion recognition skills of the participants in novel (untrained) stories.

Figure 74 presents the scores in the pre-test, post-test, and follow-up test. Participants showed difficulties to identify the relevant emotion for the social stories in the pre-test.

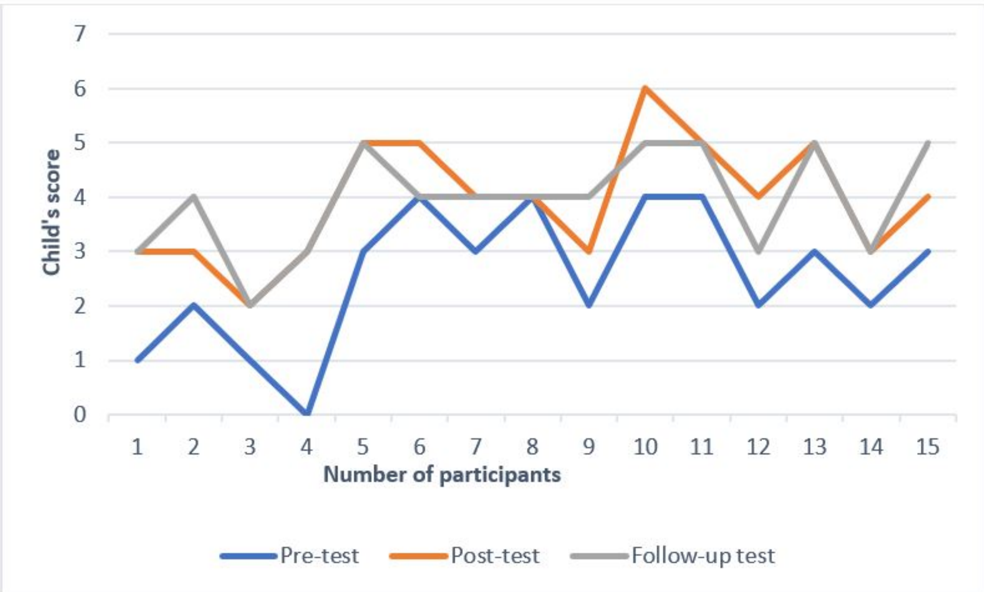


Figure 74: Comparison between the scores of the participants in the pre, post, and followup tests: phase III of the emotion recognition programme.

The mean scores of the participants were calculated as shown in Figure 75. The mean numbers of correct answers in the pre-test, post-test, and follow-up test are 2.53, 3.93, and 3.93. The participants' performance in recognising the basic six emotions from different social contexts increased by 23.3% in the untrained scenarios (from 42.2% in the pre-test to 65.5% in the post-test). Additionally, the positive learning outcomes were maintained two weeks after the training.

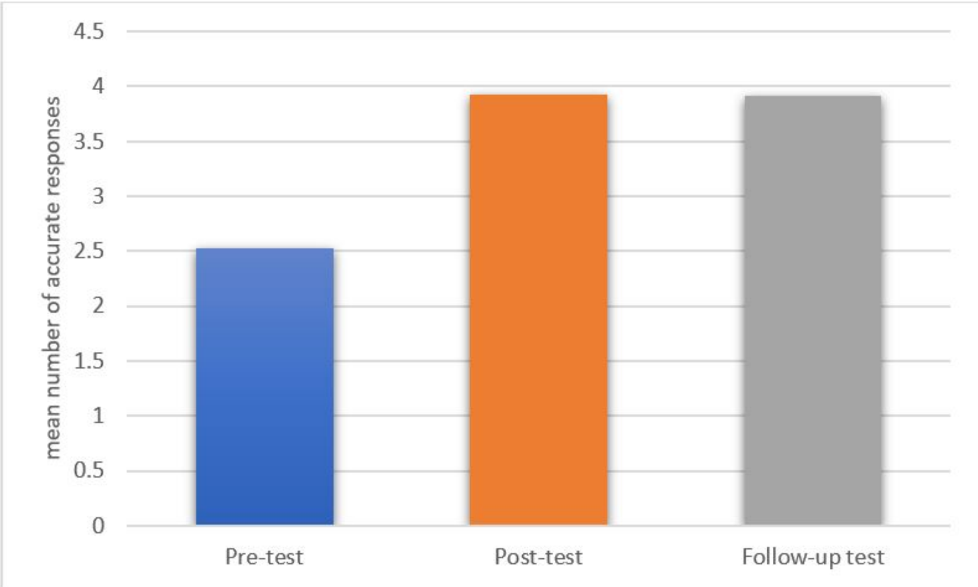


Figure 75: Mean number of accurate responses in the pre, post, and follow-up tests: Phase III of the emotion recognition programme.

Figure 76 shows the recognition rate for each of the six emotions. The recognition rates are: happiness - 75.5%, fear - 80%, sadness - 60%, anger - 42.2%, disgust - 60 %, and surprise - 28.8 %. It was observed that there was confusion in recognising HAPPY and SURPRISE emotions from the emotional-social stories. Some participants chose HAPPY instead of SURPRISE in some scenarios. Additionally, the same confusion was observed between ANGRY and SAD emotions.

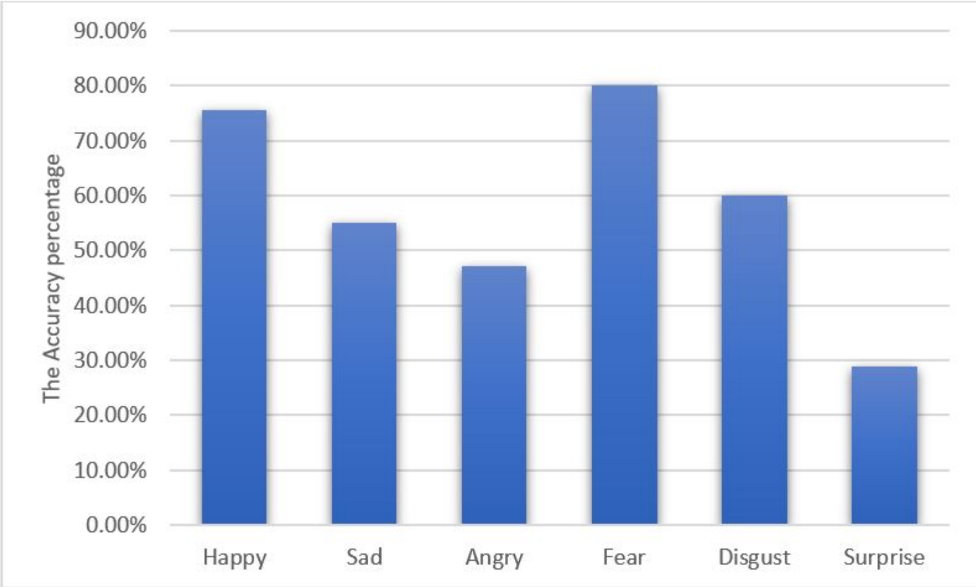


Figure 76: The accuracy percentage of the recognised emotions from social contexts in Phase III.

Table 21 illustrates the Wilcoxon statistical analysis and highlights the p-value.

Table 21: Wilcoxon test results: Emotion recognition-phase III.

	Pre-test	**Post-test**	**Follow-up test**
Sample Size	15	15	15
Mean	2.53	3.93	3.93
Standard Deviation	1.25	1.10	0.96
P-Value (pre-test and post-test)	P<0.001		
P-Value (post-test and follow-up test)		P=1.00, (p>0.05)	

- Q1: Is there a significant improvement in the participants' emotion recognition skills from social stories after the intervention?

  Two hypotheses derive from this research question.

  - H0: There is no significant improvement in the emotion recognition skills of the participants from social stories after receiving the training sessions.
  - H1: There is a significant improvement in the emotion recognition skills of the participants from social stories after receiving the training sessions.

A significant difference was found between the participants' performance in the pre-test and the post-test ($P<0.001$). Thus, hypothesis (H0) was rejected, and hypothesis (H1) was accepted. The results suggest that the Jammo VRobot helped the participants understand the character's perspective in the story.

6.4.3 Quantitative Questionnaires

As mentioned earlier, the teacher and parents were asked to complete questionnaires pre and post the intervention. The pre-questionnaire aimed to assess the target skills of the participants (emotions, imitation, and gestures) before starting the training sessions. The post-questionnaire aimed to measure the improvement in the participants' target skills after receiving the training and understand the parents/teacher opinions in the protocol and the scenarios used.

6.4.3.1 Pre and Post Questionnaires

The pre-questionnaire consists of 16 items, including three scales; 3 items for responding to interactions, 8 items for emotion recognition and expression (affective understanding), and 5 items for gesture and play skills.

The post-questionnaire contains the same 16 items in the pre-questionnaire with extra 12 items. The extra 12 items aimed to indicate the opinions of the teacher and parents about the intervention their children received. The post-questionnaire was split into two parts as some items are incomparable to the items in the pre-questionnaire. Table 6.4.3.1 illustrates the results of pre and post questionnaires.

Table 22: The results of the pre and post questionnaires.

item	Pre-questionnaire				Post-questionnaire		
	Never	Sometimes	Often	Always	Sometimes	Often	Always
Is aware of others' thoughts and feelings.		10 (66.7%)	2 (13.3 %)	3 (20%)	6 (40%)	4 (26.7%)	5 (33.3%)
Is able to correctly identify other's feelings from their facial expressions, voice tone and body postures.		11 (73.3%)	3 (20%)	1 (6.7%)	4 (26.7%)	6 (40%)	5 (33.3%)
Is aware of his/her own thoughts and feelings		7 (50%)	5 (33.3%)	3 (21.4%)	5 (33.3%)	4 (26.7%)	6 (40%)
Understands other's basic emotions. Ex: happy, sad, etc		11 (73.3%)	2 (13.3%)	2 (13.3%)	3 (20%)	6 (40%)	6 (40%)
Understands the use of basic emotions.	2 (14.3%)	9 (60%)	2 (14.3%)	2 (14.3%)	5 (33.3%)	4 (26.7%)	6 (40%)
Uses gestures to express emotions.	4 (26.7%)	8 (53.3 %)	3 (20%)		6 (40%)	4 (26.7%)	5 (33.3%)
Uses conventional facial expressions to express his/her feelings. Ex: raised eyebrows to express surprise.	3 (20%)	7 (46.7%)	5 (33.3%)		7 (46.7%)	3 (20%)	5 (33.3%)
Understands emotions through social situations.	5 (33.3%)	7 (46.7%)	3 (20%)		8 (53.3%)	3 (20%)	4 (26.7%)
Uses a wide range of gestures to communicate. Ex: right/left palm waving (Hello)		10 (66.7%)	5 (33.3%)		2 (13.3%)	8 (53.3%)	5 (33.3%)

item	Pre-questionnaire				Post-questionnaire		
	Never	Sometimes	Often	Always	Sometimes	Often	Always
Understands other's gestures.		11 (73.3%)	4 (26.7%)		3 (20%)	8 (53.3%)	4 (26.7%)
Imitates other's actions/gestures.	2 (13.3%)	10 (66.7%)	3 (20%)		2 (13.3%)	5(33.3%)	8 (53.3%)
Uses appropriate gestures in social contexts.	3 (20%)	9 (60%)	3 (20%)		3 (20%)	8(53.3%)	4 (26.7%)
Understands gestures in social situations.	4 (26.7%)	11 (73.3%)			5 (33.3%)	7 (46.7%)	3 (20%)
Responds to questions.		8 (53.3%)	7 (46.7%)		3 (20%)	9 (60%)	3 (20%)
Follow instructions.	2 (13.3%)	8 (53.3%)	5 (33.3%)		3 (20%)	11 (73.3%)	1 (6.7%)
Pay attention to the instructions.	1 (6.7%)	9 (60%)	5 (33.3%)		1 (6.7%)	13 (86.7%)	1 (6.7%)

The mean scores for the pre and post questionnaires highlight that the participants show movement from the baseline (Pre-intervention) to outcome (Post-intervention) measures as rated by the teacher and parents as shown in Table 23.

As per the parents/teacher questionnaire, the participants showed improvement in their emotion recognition and expression skills (affective understanding) from pre (M= 18.2, SD=5.32) to post (M=23.9, SD=6.32) intervention; $P<0.001$. Regarding the gesture/play skills of the participants, the difference between the mean score in the pre and post-intervention is 5.2; from pre (M=10.4, SD=2.20) to post (M=15.6, SD=3.18) intervention, and the analysis revealed significant differences as p-value is less than 0.05 ($P<0.05$). Furthermore, the children showed improvement in responding to interactions skills from pre (M=6.93, SD=1.67) to post (M=8.86, SD=1.36) with P-value less than 0.05.

Table 23: Mean and standard deviation (SD) scores for the pre-questionnaire (Baseline) and post-questionnaire (Outcome).

Scale	**Pre** Mean (SD)	**Post** Mean (SD)	**P-value** (Wilcoxon test)
Emotion recognition/expression (Affective understanding)	18.2 (5.32)	23.9 (6.32)	$P<0.001$
Gesture Play skills	10.4 (2.20)	15.6 (3.18)	$P<0.05$
Responding to interactions	6.93 (1.67)	8.86 (1.36)	$P<0.05$

Table 24 shows the acceptance of parents and teacher and their satisfaction about the training programme their children received. The post-questionnaire was completed by the teacher and parents shows that most of them found the training programme used effective and helped to improve the target skills of their children.

Table 24: Parents/teacher satisfaction about the training programme their children received.

Item	**Never**	**Sometimes**	**Often**	**Always**
The child's emotion recognition skill improved over the sessions.		1 (6.7%)	7 (46.7%)	7 (46.7%)
The emotion recognition scenarios used were appropriate to train the child emotion ability.			5 (33.3%)	10 (66.7%)
I am satisfied with the emotion recognition scenarios that my child practised.			4 (26.7%)	11 (73.3%)
I would be interested in continuing using these emotion recognition scenarios to train my child to emotion ability.		1 (6.7%)	3 (20%)	11 (73.3%)
The child can generalise emotion recognition skill that has been learnt in this study to daily life.		6 (40%)	6 (40%)	3 (20%)
The child's imitation skill improved over the sessions.			7 (46.7%)	8 (53.3%)
The child's play skills improved over the sessions.		1 (6.7%)	7 (46.7%)	7 (46.7%)
The child starts using the taught gestures in daily life.		3 (20%)	7 (46.7%)	5 (33.3%)
The child's gesture recognition skill improved over the sessions.		2 (13.3%)	7 (46.7%)	6 (40%)
The intransitive gesture scenarios used was appropriate to teach the child gesture skill.			6 (40%)	9 (60%)
I am satisfied with the gesture scenarios that my child practised.			6 (40%)	9 (60%)
I would be interested in continuing using these gesture scenarios to teach my child to more gestures.			6 (40%)	9 (60%)

6.4.3.2 System Usability Scale Results

The teacher and parents were asked to complete the System Usability Scale (SUS) questionnaire to measure usability and their satisfaction with the Jammo-VRobot environment. A teacher and 11 parents completed the SUS questionnaire at the end of the intervention sessions as shown in Table 25. The mean score on the SUS was 73.75 (SD=7.34), which corresponds to a score of "Good" (UIUXTrend, 2021). Figure 77 shows that 59% of the participants found the Jammo-VRobot tool "Good", 33% found it "Okay", and 8% found it "Excellent".

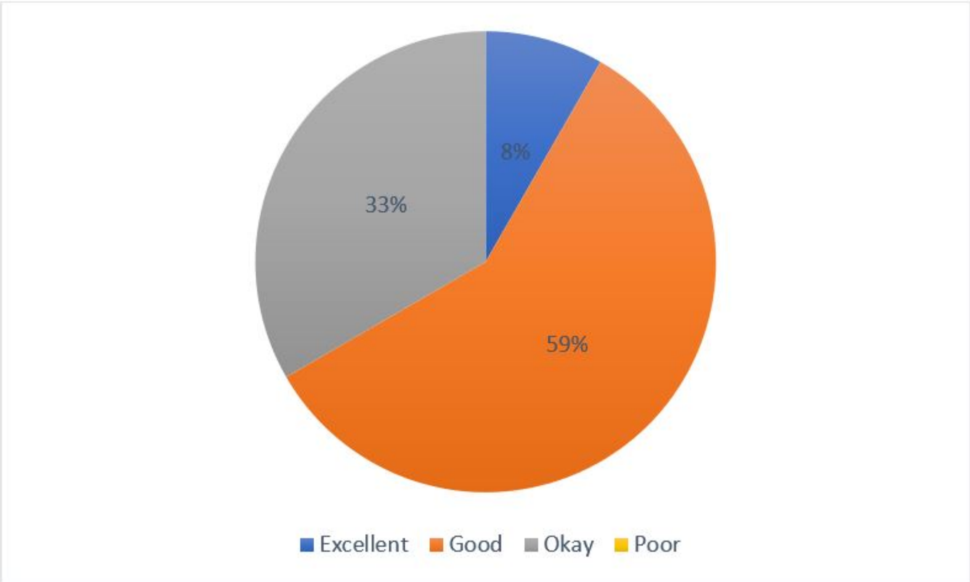


Figure 77: System usability scale chart.

Ten (83.3%) participants agreed or strongly agreed with the statements: "I think that I would like to use this tool frequently", "I found the various functions in this tool were well-integrated", "I would imagine that most people would learn to use this tool very quickly", and "I felt very confident using the tool". All the participants disagreed or strongly disagreed with the statements:" I found the tool very awkward to use" and "I thought there was too much inconsistency in this too". All the participants agreed or strongly agreed with the statement: "I thought the tool was easy to use". Ten participants (83.3%) disagreed or strongly disagreed with the statement: "I think that I would need the support of a technical person to be able to use this tool".

All the participants except one disagreed or strongly disagreed with the statement: "I found the tool unnecessarily complex". Regarding this statement: "I needed to learn a lot of things before I could get going with this tool", three participants (25%) disagreed, four (33.3%) were neutral, and five (41.7%) agreed.

Table 25: System Usability Scale (SUS) results.

Item	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I think that I would like to use this tool frequently			2 (16.7%)	9 (75%)	1 (8.3%)
I found the tool unnecessarily complex	2 (16.7%)	9 (75%)			1 (8.3%)
I thought the tool was easy to use				10 (83.3%)	2 (16.7%)
I think that I would need the support of a technical person to be able to use this tool	2 (16.7%)	8 (66.7%)	2 (16.7%)		
I found the various functions in this tool were well integrated			2 (16.7%)	7 (58.3%)	3 (25%)
I thought there was too much inconsistency in this tool	3 (25%)	9 (75%)			
I would imagine that most people would learn to use this tool very quickly			2 (16.7%)	7 (58.3%)	3 (25%)
I found the tool very awkward to use	2 (16.7%)	10 (83.3%)			
I felt very confident using the tool			2 (16.7%)	9 (75%)	1 (8.3%)
I needed to learn a lot of things before I could get going with this tool		3 (25%)	4 (33.3%)	5 (41.7%)	

These results provide evidence of the users' satisfaction regarding the usability of the Jammo-VRobot tool. Once everything was explained to the parents and teacher on how to use the Jammo-VRobot tool and navigate from one scene to another, they found it easy to use. These results reveal that the Jammo-VRobot tool is an easy tool to use without technicians or professionals.

6.4.4 Qualitative Results: Observational Data

The researcher and parents completed the observation sheet during the intervention sessions that records the child's behaviours and skills while interacting with the Jammo VRobot and exploring the training experience. The results from the observation were satisfactory. The participants' paid attention and followed the instructions given by the Jammo VRobot. It was also observed that the participants were imitating Jammo VRobot actions and words without asking them to imitate. This observation reveals that the Jammo VRobot kept the participants engaged and motivated in the intervention sessions. It was also observed that the Jammo VRobot did not bring any negative emotions (boredom, anger) during the intervention sessions. Furthermore, one of the promising observations was that the Jammo VRobot managed to reduce the repetitive behaviours of one child during the intervention sessions. In the on-site sessions, it was noticed that a child was repeating words and actions out of context; throughout the sessions, it was observed that the child's repetitive behaviours were gradually decreased. These findings indicate that the Jammo VRobot has positive effects on the participants and captures their attention.

There were many positive comments and observation notes about the child's behaviours towards the Jammo VRobot. Here are some comments and observations:

> "My child got excited towards the robot and wants to look at it all the time."
>
> "He always responds to the robot questions."
>
> "Sometimes looks away while I am talking to him, but never looks away while the robot is talking."
>
> "In the training, she imitates the robot and says the name of the gesture after it."
>
> "Explains what the robot physically does, like for ANGRY: crossed its arms and sad."
>
> "Gets excited for the training."
>
> "He keeps copying the robot's actions."
>
> "At the beginning of the session, she was doing a lot of repetitive behaviours. However, throughout the sessions, it was observed that the repetitive behaviours decreased."
>
> "The robot animation managed to grab her attention."
>
> "She loved the stories and listened very well to the instructions."
>
> "He gets really enthusiastic when I tell him that we are going to start the game."
>
> "Gets very excited when he saw the cartoon characters "Inside out cartoon" in the training."

6.4.4.1 Generalisation

Generalisation is the ability to transfer the learnt skills within a specific context to other settings and with different people. Measuring generalisation is necessary for improving the effectiveness of any social skill intervention. Children with ASD are having difficulty transferring skills to new situations and maintaining skills for longer periods (Carruthers et al., 2020; Shindorf, 2016). According to Shindorf (2016), parent acceptability of social skill interventions may influence the generalisation of the learnt skills. Therefore, involving parents in the social skills interventions may improve the social skills of children with ASD, as they might implement the learnt skills at home.

After two weeks of completing the intervention programme, the generalisation of the taught skills was discussed with the parents and teacher. The teacher and parents were asked if they notice any changes in their children. They mentioned that they had seen some improvement in the participants' emotion recognition and expressions skills. They also pointed out some enhancements in the use and recognition of gestures and imitation skills. Most of the parents reported significant improvement in their children's imitation skills. Here are some of their feedback:

> "I noticed a difference with my child ability to recognise expressions of emotions."
>
> "I have noticed an improvement in my child imitation skills. He starts to imitate my actions and my words."
>
> "He started expressing his feelings more. Sometimes he says I feel blue."
>
> "She uses more gestures now in the daily life."
>
> "I can see that his imitation skills improved as he started to copy my actions more."
>
> "I definitely see progress in my child's emotion recognition and expression skills."
>
> "He understands more gestures now than before."

The teacher was asked if he sees any differences in the participants' performance and progress between the intervention programme and the traditional sessions he usually conducts.

> "Since the robot is attractive and the animations grab their interest, the participants pay more attention and learn faster compared to the traditional session. Additionally, the consistency of the scenarios and the virtual robot behaviour is a great advantage."

The teacher feedback is in line with the educators' opinions in Alcorn et al. (2019) study. Educators emphasise that children with ASD might learn more easily and faster as well as feel more comfortable with robots than when learning with a person.

6.5 Conclusion

This chapter has described the experimental set-up, data collection methods, and the results of the conducted sessions. After designing and implementing the social skill training programme that targets three skills, the evaluation process took place. Recruiting participants was challenging due to the nature of the study, the target age group, and obtaining the parent's consent. A primary challenge was the relatively small sample size (N=15). Nonetheless, the sample was larger than that found in some state-of-the-art studies. Conducting this study during the pandemic circumstances (Covid-19) was the most significant challenge that affects the sample size. The evaluation process was conducted in Egypt. The number of assistive technology interventions for training children with ASD in Egypt is limited. Additionally, it is essential to emphasise that social robot interventions were not applied in Egypt before. Thus, the Jammo-VRobot environment provides an excellent opportunity for Egyptian parents and teachers to train their children's social skills.

Fifteen children with HFA participated in this study online (11 children) and on-site (4 children). The participants received 24 sessions, with two sessions per week. The participants were taught to recognise the basic six emotions and 11 intransitive gestures (phase I), to imitate and express these emotions and gestures (phase II), and to express and demonstrate them in social situations (phase III).

The evaluation process was divided into three phases: pre-training, training, and post-training. Data was collected through questionnaires and live observation. The data collected from the live observation was statistically analysed to evaluate the significant differences between the participants' scores in the pre-tests and post-tests. Significant differences across all the phases in each skill training programme were found between the pre-test, post-test, and follow-up test. It was observed that the girls performed better than the boys in most of the phases. In the emotion recognition and expression training programme, the girls' performance increased by 47% in phase I, 20% in phase II, and 24% in phase III, while the boys' performance increased by 43% in phase I, 31% in phase II, and 24% in phase III. In the intransitive gesture training programme, the girls' performance increased by 44% in phase I, 25% in phase II, and 16% in phase III, while boys improved by 39% in phase I, 13% in phase II, and 15% in phase III. These results might suggest that the girls' attention span during the intervention sessions was higher than boys. The results also indicate that girls did better using the Jammo-VRobot tool. In the literature, to our knowledge, none of the relevant studies has analysed the results based on gender; therefore, we can not compare our findings. Additionally, it was observed that the scores of the young participants in the pre-tests were less than the scores of the rest of the participants (ex: participants 8 and 9 aged 5 and 4 years old). These low scores might be due to their young age, gender (both participants are boys), and the novelty of the training. For example, the average scores of the participants aged between 4 - 6 years in phase I of the intransitive gesture training programme are 3.2. It was also noticed that the improvements were found to be affected by the age of the participants. The Jammo-VRobot tool is more effective for younger age (4-8 years). This finding emphasises the importance of early intervention for training the social skills of children with ASD. It might also suggest that younger participants engaged more with the Jammo-VRobot tool due to the use of the robot animations. Moreover, the teacher's feedback emphasises that animations and the designed virtual environments made the learning process more enjoyable than the traditional way.

The teacher and parents were asked to complete the pre and post questionnaires to assess the effectiveness of the Jammo-VRobot environment. Pre and post questionnaires suggest an improvement in the participants' learnt skills and potential for generalising those skills in daily life. It is worth pointing out that parents and teachers have more contact with their children and are more able to observe the changes in their skills; however, they might be biased towards their children. Furthermore, previous studies on social robot and VEs interventions have assessed the generalisation either through observation in a natural setting or through parents' and teachers' reports. However, these studies did not provide statistical results regarding the generalisation makes it hard to compare our findings in generalisation with the state-of-the-art.

At the beginning of the experimental sessions, there were some challenges. The intention was to conduct these experimental sessions in private special needs schools in Egypt; thus, English was not a problem. Due to the precautions measures, the arrangements that were made with these schools became no longer available. Therefore, an Arabic version of the Jammo-VRobot environment was developed as requested by some parents online and local specialised centre to increase the environment's potential reach.

7 Discussion and Conclusion

7.1 Discussion

The primary aim of this study was to assess the effectiveness of the Jammo-VRobot environment in training children with HFA on emotion recognition and expression, gesture recognition and production, and imitation skills. In addition, the main objective was to provide a cost-effective and user-friendly tool for parents and practitioners to use.

Regarding the research questions, the following implications were found:

Could the virtual robot be as effective as the physical robot in enhancing the social skills of children with ASD?

This intervention study presents encouraging results showing that the Jammo-VRobot environment that utilised virtual robot (Jammo VRobot) helps training and enhancing the emotion recognition and expression, gesture recognition and production, and imitation skills of children with HFA.

By comparing the results obtained from the Jammo-VRobot social skill training programme environment with the state-of-the-art, comparable results were produced at a much lower cost and without needing technicians or professionals to set up the environment. Additionally, the Jammo-VRobot environment made it possible to explore the effect of such intervention in countries like Egypt, where no social robot interventions have been conducted before. Furthermore, providing the Jammo-VRobot tool as open-source tool will allow practitioners to expand the designed social skill training programme to include more social skills.

Intransitive Gesture Training Programme

As mentioned earlier, the intransitive gesture training protocol was adapted and modified from So et al. (2018b) that used the same protocol with a physical (NAO) robot. The authors recruited 20 children with ASD, and they were taught to recognise and produce 20 gestures. Table 26 shows a comparison with the state-of-the-art regarding the mean of accurate responses in the pre-test, post-test, and follow-up test in the three phases. Considering the difference in sample size and the number of gestures between our study (N=15, gestures=11) and So et al. (2018b) study (N=20, gestures=20), the comparison shows comparable findings.

Table 26: Comparison with the state-of-the-art on the intransitive gesture training programme.

	Current study			State-of-the-art (So et al., 2018b)		
	Pre-test	Post-test	Follow-up test	Pre-test	Post-test	Follow-up test
Phase I	4	6.8	8.4	12.5	16.5	16.3
Phase II	6.46	8.46	8.33	9	11	11.8
Phase III	2.4	4.3	4.1	11.3	13	12.3

Results produced from So et al. (2018b) show that the participants' gestures recognition skills increased by 20% in comparison between the pre-test and the post-test (from 62.5% in the pre-test to 82.5% in the post-test). Our results show slightly better in the participants' gesture recognition skills in phase I. 25.5% increase from 36.3% in the pre-test to 61.8% in the post-test. In phase II, the gesture expression skill of the participants increased by 18.2% from the pre-test (58.7%) to the post-test (76.9%) in our training programme, while So et al. (2018b) show 10% increase in their participants' gesture expression skills (from 45% in the pre-test to 55% in the post-test). The results obtained from phase III of the study carried by So et al. (2018b) show 8.3% in the participants' skills, while our intervention shows better results by a 16.9% increase in the participants' skills.

Emotion Recognition and Expression Training Programme

Soares et al. (2019) utilised Zeno robot to teach 15 children with ASD to recognise and express the basic five emotions: happiness, fear, sadness, anger, and surprise.

The results obtained from the Jammo-VRobot emotion recognition training programme were compared to the results obtained by (Soares et al., 2019) as shown in Table 27.

Table 27: Comparison with the state-of-the-art on the emotion recognition training programme (percentage of performance).

	Current study		State-of-the-art (Soares et al., 2019)	
	Pre-test	Post-test	Pre-test	Post-test
Phase I	31.1%	71%	50.5%	73.4%
Phase II	52%	78.8%	67.5%	83.7%
Phase III	42.2%	65.5%	62.7%	82.3%

In the emotion recognition phase, Soares et al. (2019) results indicate that there was a 23% increase in their participants' emotion recognition skills. In contrast, our results show a 40% increase in the participants' performance in recognising the basic six emotions. In the expression phase, the performance of our participants increased by 26.8%. On the other hand, the other participants show a 16.2% increase in emotion expression skills. Finally, the results obtained from recognising the relevant emotion from social contexts show a 23.3% increase in our participants' performance and 19.5% in their study.

Soares et al. (2021) conducted a recent preliminary study that utilised a virtual robot (HiZeca) for training the emotion recognition and expression skills of children with ASD. HiZeca interacts with children through three levels: imitate, recognise, and identify the relevant emotion. Two children with ASD (aged 6-8 years) participated in the study to verify the acceptance of the developed tool. The participants' feedback was positive regarding the acceptability of the tool. The initial results show that DISGUST and SURPRISE were the most difficult emotions to recognise, followed by the FEAR emotion. This finding is in line with our findings that DISGUST and SURPRISE emotions are the most difficult emotions to recognise. This study is still in its early stages and was published after the Jammo-VRobot environment experimental sessions were conducted, and their results were published.

The results obtained across all three phases for each skill emphasise that Jammo VRobot resulted in slightly better results than the physical robot in improving the target skills of children with HFA. According to the literature, sometimes children with ASD focus on the physical aspects of the robot itself and want to touch it rather than focusing on the training scenarios (Soares et al., 2019). Additionally, some robots' sounds produced from the motors while moving could distract children with ASD, as mentioned by some professionals and specialists (Huijnen et al., 2017). Therefore, this might suggest why a virtual robot might be an effective or a good teaching tool and increase the participants' focus skills during the learning process. The animation and cooperating cartoon characters in the learning process increased the participants' excitement during the intervention session and facilitated the learning process.

Additionally, there is a need for cost-effective training tools for children with ASD that can be used at home, in the classroom, or as a supplementary method with traditional training in therapy sessions. Therefore, a virtual robot is an excellent opportunity if a real (physical) robot is not available. Furthermore, providing a widely available tool that can be easily used by parents and practitioners either at home or school can result in repetitive training, which is the key to success in training children with ASD.

Thus, the results reveal that the virtual robot is as effective as the real (physical) robot in enhancing the social skills of children with HFA. These results support previous studies demonstrating the usefulness of robot interventions with children with ASD and provide new evidence to the usefulness of virtual robots to train and improve the social skills of children with ASD. Moreover, it emphasises that the Jammo-VRobot could overcome some of the limitations of physical social robot interventions, their cost and limited availability and became a widespread tool that parents and teachers can easily use to train the social skills of children with ASD.

To what extent does the Jammo-VRobot environment help children with HFA recognise and comprehend social emotions, improve their imitation skills, and recognise and produce gestures?

Significant differences were found between pre-test, post-test, and follow-up test in the three phases of each training programme among 15 children with HFA. Regarding the intransitive gesture training programme, the number of successful answers in the three phases increased after receiving the four training sessions as shown in Figure 78. As mentioned in the results section, the participants' performance in producing the 11 taught gestures in the pre-test (phase II) was higher than their performance in recognising the gestures (phase I). This result shows that the learning of gesture recognition is contributed to the learning of gesture production due to the spontaneous imitation that occurred in the training session of phase I. It was observed that the participants were imitating the Jammo VRobot without asking them to imitate. This finding is in line with previous studies that indicate that it is crucial to teach children with ASD to recognise meaningful gestures before asking them to imitate and produce these gestures (So et al., 2018b). Our findings confirm that imitation ability is essential to the development of language skills of children with ASD. It was observed that the participants were verbally imitating the name of the gestures after Jammo VRobot in the training sessions. Additionally, imitation is proof of the participant's engagement with the Jammo-VRobot environment. Although the participants' scores in the storytelling tasks (Phase III) were lower than the other tasks (phase I and phase II), their scores in the untrained scenarios (post-test and follow-up test) were higher than their scores in the pre-test. More importantly, the positive learning outcomes were maintained two week after the training in the three phases when a follow-up test was given.

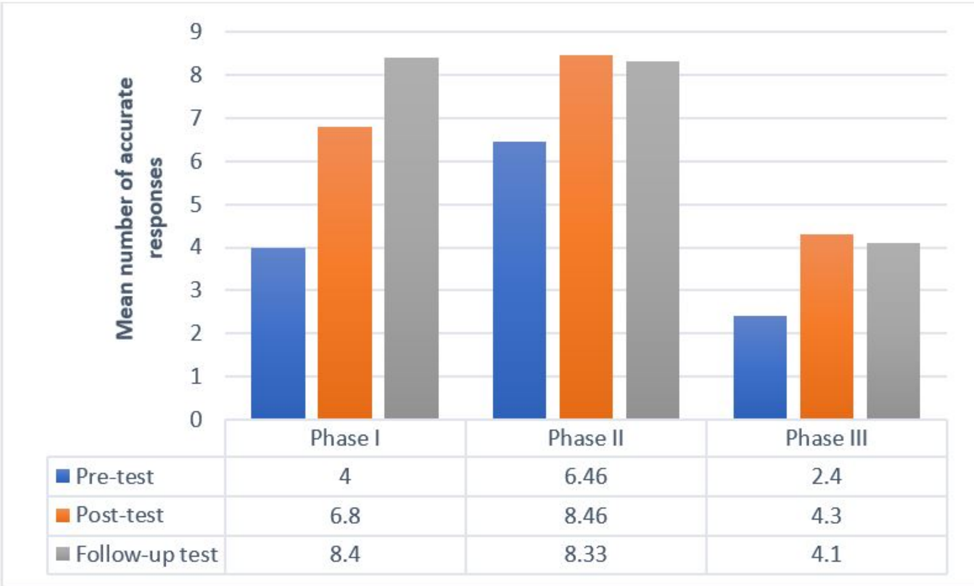


Figure 78: Mean number of correct responses in the intransitive gesture training programme: phase I, phase II, and phase III.

Figure 79 shows the mean scores of the participants in each pre-test, post-test, and follow-up test in the three phases of the emotion recognition and expression training programme. The participants kept improving along with the sessions. Differences were found by comparing the scores of the pre-tests and post-tests in each phase. The statistical analysis highlighted a significant improvement in the participants' emotion recognition and expression skills (phase I <0.001; phase II <0.001; phase III <0.001). Additionally, the Jammo VRobot taught the participants to associate a colour to each emotion.

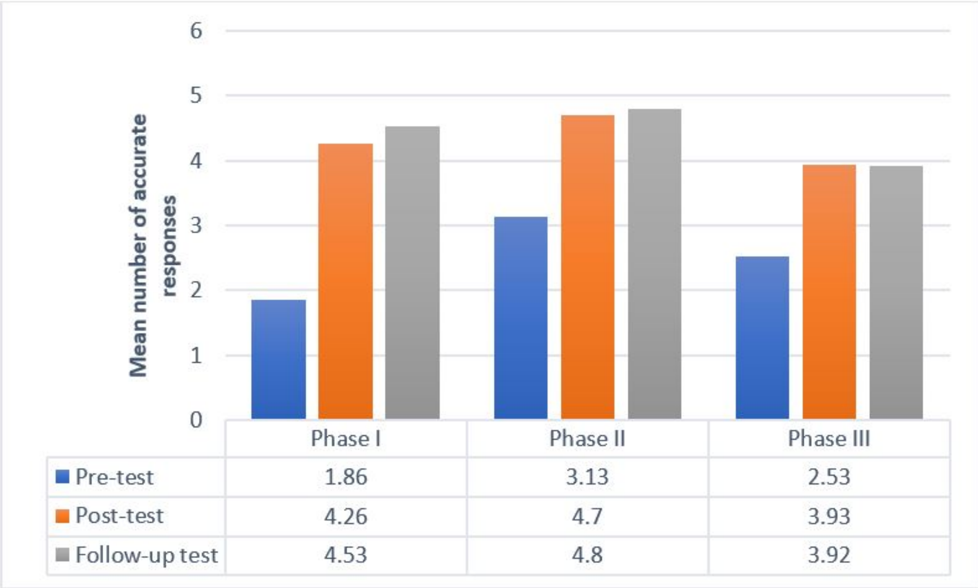


Figure 79: Mean number of correct responses in the emotion training programme: phase I, phase II, and phase III.

As documented in Chapter 6, some emotions were easier than others to recognise and express. HAPPY and SAD emotions were the easiest to recognise and express among the 15 participants in the three phases as shown in Table 28. DISGUST and SURPRISE emotions were the hardest to recognise and express in phase I and phase II, which indicates the importance of facial expressions to express such emotions. In phase III, the participants found it hard to differentiate between HAPPY and SURPRISE emotions from the social stories. Most of the participants chose HAPPY instead of SURPRISE. The same confusion was observed between ANGRY and SAD emotions.

Table 28: The children's mean score, in the three phases, for each emotion.

Emotion	Phase I	Phase II	Phase III
Happy	0.86	1	0.75
Sad	0.82	1	0.55
Fear	0.51	0.66	0.8
Angry	0.8	0.8	0.47
Disgust	0.26	0.33	0.6
Surprise	0.33	0.42	0.28

The outcomes from the parents and teacher questionnaires suggest an improvement in the participants' taught skills. Figure 80 illustrates the difference in means between the baseline (pre) and outcomes (post) measures in three scales; affective understanding, appropriate gesture responses, and responding to interactions.

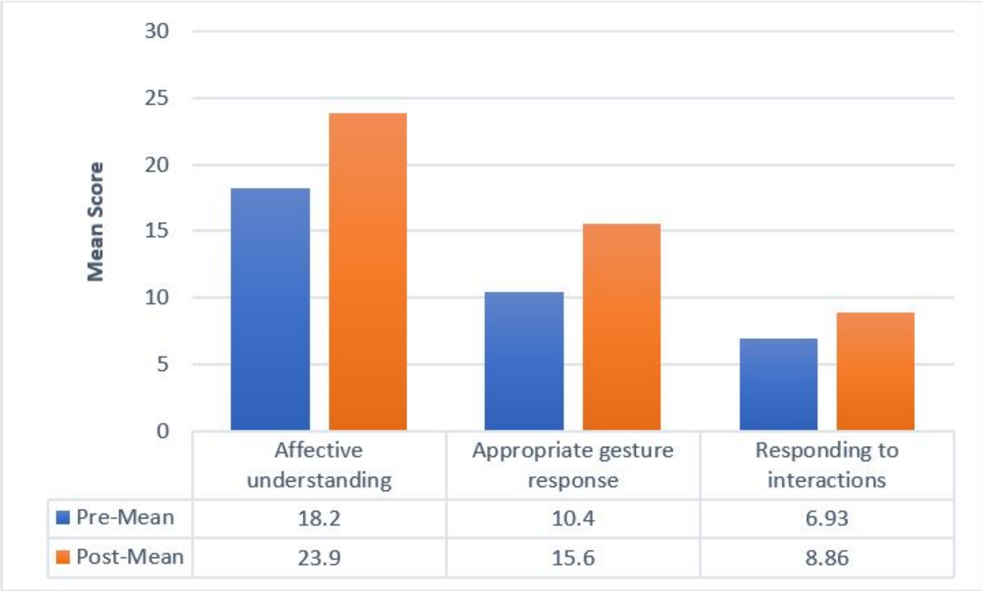


Figure 80: Mean score for the pre and post questionnaires in the three measured scales.

The observation of the intervention sessions brought valuable information regarding the participants' behaviours and skills during the sessions. It was observed that the children got engaged with the Jammo VRobot and imitated its actions and words spontaneously. Additionally, it was noticed that the Jammo VRobot did not bring any negative emotions, and most of the children were excited during the intervention sessions.

However, this study was not focused on repetitive behaviours and joint attention. It was observed that the Jammo VRobot increased the participants' attention and reduced the repetitive behaviours of one child in the sessions. It is important to emphasise that the presence of the teacher and parents in the intervention sessions ensured the smooth conduct of the sessions. They know the participants' habits and what could make them anxious and help them work better.

Furthermore, the educator who took place in this study was highly satisfied with the intervention programme. The consistency in the training programme and how the Jammo VRobot expresses and produces emotions and gestures were standardised and constantly repeated the same way, unlike a human, made the learning process faster than the traditional training. Additionally, the teacher's statement indicates that the Jammo VRobot positively influenced the participants' development.

These results conclude that the participants' skills progressively improved over the training sessions.

Can the Jammo-VRobot environment help children with ASD to generalise the skills they learnt in the intervention session to their daily life?

As mentioned in Chapter 6, the teacher and parents were asked about transferring the taught skills in the intervention sessions to daily life. 80% of parents and teacher stated that their children either often or always start using the taught gestures in daily life. 93.4% agreed or strongly agreed that the emotion recognition skills of their children improved over the sessions. The teacher and parents agreed or strongly agreed that the participants' imitation skills are improved.

Although achieving generalising is challenging for children with ASD, the results indicate the potential for generalising the learnt skills.

7.2 Limitations

There are some limitations in this research that future research might addresses. A primary challenge was conducting this study during pandemic circumstances (Covid-19). The arrangements were made with specialised schools, and centres in Egypt, and these became no longer viable due to the precautionary measures. The pandemic circumstances have limited the number of on-site participants and data collection methods. Increasing the number of on-site participants would help draw better insights into the effectiveness of the Jammo-VRobot environment. Video observation and multiple perspective questionnaires would provide additional information regarding the participants' skills and behaviours during the real-time interaction with the environment that could be missed during the live observation and avoid the concerns regarding the parents being biased. Due to the previous limitation, the sample size has been affected. A relatively small sample size has led to the absence of a control group. Although the results reveal the effectiveness of the Jammo-VRobot environment in teaching the target skills to the participants, a control group would help to generalise these findings to other children with HFA.

Another limitation is the lack of facial expressions in the Jammo VRobot. Nonetheless, the Jammo VRobot is still capable of expressing emotions through other means, including gestures, colours, and eye shape and incorporating popular cartoon characters was an additional resource of information. The Jammo-VRobot has been a successful tool in training the emotion recognition and expressions skills of children with HFA. Further investigation might involve more sophisticated (expressive) robots.

7.3 Future Work

Future research with a larger sample of children with ASD will help to generalise the obtained results. Furthermore, it is essential to expand the current research and adjust the developed training programme to include more severe forms of autism with different developmental skills; this would allow children with low-functioning autism (LFA) to use it. The participants in this study were children with HFA, but this type of intervention may be effective for typically developing (Non-ASD) children who have difficulty in their social skills. Additionally, the research could be expanded to look at any other coexisting disorders besides ASD and observe any individual differences among the participants. This study focused on three social skills: imitation, emotion recognition and expressions, and intransitive gesture. The Jammo-VRobot training programme can be expanded to other social skills. Finally, recruiting participants from different countries will help evaluate the cultural differences and how children from different countries respond to such interventions.

7.4 Conclusion

The rise in the number of children diagnosed with ASD has led to a strong need for early interventions. Traditional social skills training interventions and assistive technology for cognition were investigated. Thus, virtual environments (VEs) and physical social robot interventions were the focus of this research. Despite the positive outcomes from virtual environments and physical social robots interventions, their widespread use is limited. Therefore, this research investigated an alternative approach to benefit from the advantages of VEs and physical social robot interventions while addressing some of their limitations.

The first contribution of this research was accomplished by developing a virtual environment (Jammo-VRobot environment) that combines VEs and a social robot as a hybrid approach to train the social skills of children with ASD and address some limitations of the previous VEs and physical social robot interventions. The Jammo-VRobot environment is a virtual desktop environment (non-immersive) that employs a 3D robot (Jammo VRobot) to interact with children through a social skill training programme. The social skill training programme targets three social skills: imitation, emotion recognition and expression, and intransitive gestures. This research was built on similar training programmes that have used physical robots to train the social skills of children with ASD. However, it differs from the previous studies as it combines different training programmes in one main training programme.

The Jammo-VRobot environment combines some elements from successfully utilised methods in VEs and social robot interventions. The Jammo-VRobot environment incorporates popular cartoon characters that represent emotions as an additional source of information, that result in facilitating the learning process and making the intervention sessions more fun. Different virtual environments were designed to illustrate the context of social stories used, as children with ASD pay more attention to visual cues. Additionally, representing social stories as animations motivates children with ASD and makes them engaged with the intervention. Building on the knowledge gathered from the physical social robot interventions: the Jammo-VRobot environment utilises a virtual humanoid robot. Professionals emphasise that children with ASD engage more with human-like embodiment than animal-like embodiment. Providing feedback was proven to keep the children motivated during the sessions. Jammo-VRobot provides immediate feedback regarding the children's actions by either complimenting or encouraging them to try another one.

In Egypt, caring for children with ASD is daunting and overwhelming for mothers. The services for children with disabilities in general and specifically for children with ASD are minimal or even absent. Children with ASD face a lack of special education and access to equal opportunity because they either drop out of mainstream schools or parents cannot afford the expenses of private schools and centres. Additionally, the number of technology interventions for children with ASD in Egypt are limited, particularly social robot interventions. Thus, the second contribution was accomplished by providing a tool that overcomes the cost and limited availability of traditional and assistive technology interventions. Providing an Arabic version of the Jammo-VRobot tool besides the online version maximises the potential reach of this environment. It is also important to emphasise that involving parents and teachers of children with ASD in the intervention sessions is beneficial. Parent's acceptability plays an essential role in generalising the learnt skills. Their belief that the intervention is effective makes them keen to practise the learnt skills with their children at home.

Experimental sessions were conducted to evaluate the effectiveness of the Jammo-VRobot environment. The experimental sessions were conducted mostly online with 11 children with HFA and on-site with 4 children with HFA. Based on a qualitative and quantitative analysis, it can be concluded that the Jammo-VRobot environment was an effective tool in training and improving the target social skills of the participants. Finally, the third contribution was accomplished by comparing the results achieved through the interaction with Jammo VRobot with those achieved through interaction with physical robots used by the state-of-the-art. The comparison reveals that the Jammo-VRobot environment obtained comparable results in a more affordable and accessible way. It also provides new evidence of the effectiveness of virtual robots in training the social skills of children with ASD.

References

Abdelmohsen, M., 2020. Enhancing social skills of children with asd. Available at http://www.sstprogrammeforchildrenwithasd.com/. [Accessed 2 October 2020].

Abdelmohsen, M., Arafa, Y., 2021. Training social skills of children with asd through social virtual robot, in: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), IEEE. pp. 314–319.

Abirached, B., Zhang, Y., Aggarwal, J.K., Tamersoy, B., Fernandes, T., Miranda, J.C., Orvalho, V., 2011. Improving communication skills of children with asds through interaction with virtual characters, in: 2011 IEEE 1st international conference on serious games and applications for health (SeGAH), IEEE. pp. 1–4.

Akeytsu, .

Wendy Stone et al., ., 2010. Assessing children with autism spectrum disorders. Available at https://vkc.mc.vanderbilt.edu/assets/files/resources/tssamanual.pdf. [Accessed 01 March 2020].

Alcorn, A., Pain, H., Rajendran, G., Smith, T., Lemon, O., Porayska-Pomsta, K., Foster, M.E., Avramides, K., Frauenberger, C., Bernardini, S., 2011. Social communication between virtual characters and children with autism, in: international conference on artificial intelligence in education, Springer. pp. 7–14.

Alcorn, A.M., Ainger, E., Charisi, V., Mantinioti, S., Petrovic, S., Schadenberg, B.R., Tavassoli, T., Pellicano, E., 2019. Educators' views on using humanoid robots with autistic learners in special education settings in england. Frontiers in Robotics and AI 6, 107.

Alzyoudi, M., Sartawi, A., Almuhiri, O., 2015. The impact of video modelling on improving social skills in children with autism. British Journal of Special Education 42, 53–68.

Aryania, A., Aghdasi, H.S., Beccaluva, E.A., Bonarini, A., 2020. Social engagement of children with autism spectrum disorder (asd) in imitating a humanoid robot: a case study. SN Applied Sciences 2, 1–17.

Askari, F., 2018. Studying facial expression recognition and imitation ability of children with autism spectrum disorder in interaction with a social robot .

Asperger, H., Frith, U.T., 1991. 'autistic psychopathy'in childhood. .

Assistivetechnology, 2022. Assistivetechnology. Available at https://www.atia.org/home/at-resources/what-is-at/. [Accessed 30 January 2022].

Attwood, T., 2003. Is there a difference between asperger's syndrome and high functioning autism. Sacramentoasis. com .

Audacity, 2020. Free open source cross platform audio software. Available at https://www.audacityteam.org/. [Accessed 03 June 2020].

Autism Speaks, 2020. Occupational therapy (ot). Available at https://www.autismspeaks.org/occupational-therapy-ot-0#:~:text=For20people20with20autism2C20OT,person's20current20level20of20ability. [Accessed 13 August 2020].

Axelrod, S., McElrath, K.K., Wine, B., 2012. Applied behavior analysis: Autism and beyond. Behavioral Interventions 27, 1–15.

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B., 2001. Recent advances in augmented reality. IEEE computer graphics and applications 21, 34–47.

Baron-Cohen, S., 2008. Autism and Asperger syndrome. Oxford University Press.

Barry, O., 2015. A video-based virtual environment for teaching social skills to adolescents with autism: in search of generalisation. Ph.D. thesis. Cardiff University.

Bartneck, C., 2002. Integrating the occ model of emotions in embodied characters .

Bauman, M.L., Kemper, T.L., 2005. Neuroanatomic observations of the brain in autism: a review and future directions. International journal of developmental neuroscience 23, 183–187.

Beaumont, R., Sofronoff, K., 2008. A multi-component social skills intervention for children with asperger syndrome: The junior detective training program. Journal of Child Psychology and Psychiatry 49, 743–753.

Bekele, E., Zheng, Z., Swanson, A., Crittendon, J., Warren, Z., Sarkar, N., 2013. Understanding how adolescents with autism respond to facial expressions in virtual reality environments. IEEE transactions on visualization and computer graphics 19, 711–720.

Berenguer, C., Baixauli, I., Gómez, S., Andrés, M.d.E.P., De Stasio, S., 2020. Exploring the impact of augmented reality in children and adolescents with autism spectrum disorder: A systematic review. International Journal of Environmental Research and Public Health 17, 6143.

Berggren, S., Fletcher-Watson, S., Milenkovic, N., Marschik, P.B., Bölte, S., Jonsson, U., 2018. Emotion recognition training in autism spectrum disorder: A systematic review of challenges related to generalizability. Developmental neurorehabilitation 21, 141–154.

Bharatharaj, J., Huang, L., Al-Jumaily, A., Mohan, R.E., Krägeloh, C., 2017. Sociopsychological and physiological effects of a robot-assisted therapy for children with autism. International Journal of Advanced Robotic Systems 14, 1729881417736895.

Boat, T.F., Wu, J.T., 2015. Mental disorders and disabilities among low-income children .

Bohlander, A.J., Orlich, F., Varley, C.K., 2012. Social skills training for children with autism. Pediatric Clinics 59, 165–174.

Bondy, A.S., Frost, L.A., 1998. The picture exchange communication system, in: Seminars in speech and language, © 1998 by Thieme Medical Publishers, Inc.. pp. 373–389.

Bozgeyikli, L., 2016. Virtual reality serious games for individuals with autism spectrum disorder: Design considerations .

Bozgeyikli, L., Raij, A., Katkoori, S., Alqasemi, R., 2018. A survey on virtual reality for individuals with autism spectrum disorder: Design considerations. IEEE Transactions on Learning Technologies 11, 133–151.

Bradley, R., Newbutt, N., 2018. Autism and virtual reality head-mounted displays: a state of the art systematic review. Journal of Enabling Technologies .

Britto, T., Pizzolato, E., 2016. Towards web accessibility guidelines of interaction and interface design for people with autism spectrum disorder, in: ACHI 2016: the ninth international conference on advances in computer-human interactions, pp. 1–7.

Calvo, R., D'Mello, S., Gratch, J., Kappas, A., Lhommet, M., Marsella, S., 2015. Expressing emotion through posture and gesture.

Cao, H.L., Esteban, P., Bartlett, M., Baxter, P.E., Belpaeme, T., Billing, E., Cai, H., Coeckelbergh, M., Costescu, C., David, D., et al., 2019. Robot-enhanced therapy: development and validation of a supervised autonomous robotic system for autism spectrum disorders therapy. IEEE Robotics and Automation Magazine .

Cao, H.L., Simut, R.E., Desmet, N., De Beir, A., Van De Perre, G., Vanderborght, B., Vanderfaeillie, J., 2020. Robot-assisted joint attention: A comparative study between children with autism spectrum disorder and typically developing children in interaction with nao. IEEE Access .

Carlier, S., Van der Paelt, S., Ongenae, F., De Backere, F., De Turck, F., 2020. Empowering children with asd and their parents: Design of a serious game for anxiety and stress reduction. Sensors 20, 966.

Carruthers, S., Pickles, A., Slonims, V., Howlin, P., Charman, T., 2020. Beyond intervention into daily life: A systematic review of generalisation following social communication interventions for young children with autism. Autism Research 13, 506–522.

Carter, E.J., Hyde, J., Williams, D.L., Hodgins, J.K., 2016. Investigating the influence of avatar facial characteristics on the social behaviors of children with autism, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, ACM. pp. 140–151.

CDC, 2020. Centres for disease control and prevention. Available at https://www.cdc.gov/ncbddd/autism/data.html. [Accessed 01 January 2020].

Cervera, G.R., Romero, M.G.M., Mas, L.A., Delgado, F.M., 2011. Intervention models in children with autism spectrum disorders. Autism spectrum disorders-from genes to environment , 133–156.

Chang, Y.C., Locke, J., 2016. A systematic review of peer-mediated interventions for children with autism spectrum disorder. Research in autism spectrum disorders 27, 1–10.

Charlton, C.T., Kellems, R.O., Black, B., Bussey, H.C., Ferguson, R., Goncalves, B., Jensen, M., Vallejo, S., 2020. Effectiveness of avatar-delivered instruction on social initiations by children with autism spectrum disorder. Research in Autism Spectrum Disorders 71, 101494.

Chen, C.H., Lee, I.J., Lin, L.Y., 2015. Augmented reality-based self-facial modeling to promote the emotional expression and social skills of adolescents with autism spectrum disorders. Research in developmental disabilities 36, 396–403.

Chen, C.H., Lee, I.J., Lin, L.Y., 2016. 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. Computers in Human Behavior 55, 477–485.

Cheng, Y., Huang, C.L., Yang, C.S., 2015. Using a 3d immersive virtual environment system to enhance social understanding and social skills for children with autism spectrum disorders. Focus on Autism and Other Developmental Disabilities 30, 222–236.

Cohen, I., Looije, R., Neerincx, M.A., 2011. Child's recognition of emotions in robot's face and body, in: Proceedings of the 6th international conference on Human-robot interaction, pp. 123–124.

Consequentialrobotics, 2020. Miro. Available at `http://consequentialrobotics.com/miro-beta`. [Accessed 11 July 2020].

Costa, A.P., Charpiot, L., Lera, F.R., Ziafati, P., Nazarikhorram, A., Van Der Torre, L., Steffgen, G., 2018. More attention and less repetitive and stereotyped behaviors using a robot with children with autism, in: 2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), IEEE. pp. 534–539.

Costa, A.P., Steffgen, G., Lera, F.R., Nazarikhorram, A., Ziafati, P., 2017. Socially assistive robots for teaching emotional abilities to children with autism spectrum disorder, in: 3rd Workshop on Child-Robot Interaction at HRI.

Costa, S., Lehmann, H., Dautenhahn, K., Robins, B., Soares, F., 2015. Using a humanoid robot to elicit body awareness and appropriate physical interaction in children with autism. International journal of social robotics 7, 265–278.

Costa, S., Soares, F., Pereira, A.P., Santos, C., Hiolle, A., 2014. A pilot study using imitation and storytelling scenarios as activities for labelling emotions by children with autism using a humanoid robot, in: 4th International Conference on Development and Learning and on Epigenetic Robotics, IEEE. pp. 299–304.

Cunningham, A.B., 2012. Measuring change in social interaction skills of young children with autism. Journal of Autism and Developmental Disorders 42, 593–605.

Daneshvar, S.D., Charlop, M.H., Berry Malmberg, D., 2019. A treatment comparison study of a photo activity schedule and social stories for teaching social skills to children with autism spectrum disorder: brief report. Developmental neurorehabilitation 22, 209–214.

Dapretto, M., Davies, M.S., Pfeifer, J.H., Scott, A.A., Sigman, M., Bookheimer, S.Y., Iacoboni, M., 2006. Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. Nature neuroscience 9, 28.

David, D.O., Costescu, C.A., Matu, S., Szentagotai, A., Dobrean, A., 2020. Effects of a robot-enhanced intervention for children with asd on teaching turn-taking skills. Journal of Educational Computing Research 58, 29–62.

Davis, S., 2018. Robots and autism: How children with autism spectrum disorder perceive social robots .

De Winter, J.C., 2013. Using the student's t-test with extremely small sample sizes. Practical Assessment, Research, and Evaluation 18, 10.

Delano, M.E., 2007. Video modeling interventions for individuals with autism. Remedial and Special Education 28, 33–42.

DELLTechnologies, 2020. The robot enhancing how children with autism learn. Available at https://www.delltechnologies.com/en-us/perspectives/the-robot-enhancing-how-children-with-autism-learn/. [Accessed 25 July 2021].

Demetriou, E., Lampit, A., Quintana, D., Naismith, S., Song, Y., Pye, J., Hickie, I., Guastella, A., 2018. Autism spectrum disorders: a meta-analysis of executive function. Molecular psychiatry 23, 1198.

Desideri, L., Di Santantonio, A., Varrucciu, N., Bonsi, I., Di Sarro, R., 2020. Assistive technology for cognition to support executive functions in autism: A scoping review. Advances in Neurodevelopmental Disorders 4, 330–343.

Di Nuovo, A., Bamforth, J., Conti, D., Sage, K., Ibbotson, R., Clegg, J., Westaway, A., Arnold, K., 2020. An explorative study on robotics for supporting children with autism spectrum disorder during clinical procedures, in: Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, pp. 189–191.

Dickstein-Fischer, L.A., Crone-Todd, D.E., Chapman, I.M., Fathima, A.T., Fischer, G.S., 2018.

Socially assistive robots: current status and future prospects for autism interventions. Innovation and Entrepreneurship in Health 5, 15–25.

Didehbani, N., Allen, T., Kandalaft, M., Krawczyk, D., Chapman, S., 2016. Virtual reality social cognition training for children with high functioning autism. Computers in human behavior 62, 703–711.

Diehl, J.J., Schmitt, L.M., Villano, M., Crowell, C.R., 2012. The clinical use of robots for individuals with autism spectrum disorders: A critical review. Research in autism spectrum disorders 6, 249–262.

Disney, 2020. Disney characters. Available at `https://disney.co.uk/`. [Accessed 1 January 2020].

Edelson, S.M., 2016. Understanding and treating self-injurious behavior. URL: https://www. autism. com/symptoms_self-injury.–Text: electronic .

Ekman, P., 1999. Basic emotions. Handbook of cognition and emotion 98, 16.

English, B.A., Coates, A., Howard, A., 2017. Recognition of gestural behaviors expressed by humanoid robotic platforms for teaching affect recognition to children with autism-a healthy subjects pilot study, in: International Conference on Social Robotics, Springer. pp. 567–576.

Estratopoulou, M.A., Sofologi, M., 2019. Working memory abilities in children with autism spectrum disorder. European Journal of Special Education Research .

Feng, Y., Jia, Q., Wei, W., 2018. A control architecture of robot-assisted intervention for children with autism spectrum disorders. Journal of Robotics 2018.

Flippin, M., Reszka, S., Watson, L.R., 2010. Effectiveness of the picture exchange communication system (pecs) on communication and speech for children with autism spectrum disorders: A meta-analysis. American Journal of Speech-Language Pathology .

Foden, T., Anderson, C., 2011. Social skills interventions: Getting to the core of autism. Interactive Autism Network .

Frith, U., 2001. Mind blindness and the brain in autism. Neuron 32, 969–979.

Ganz, J.B., Simpson, R.L., 2004. Effects on communicative requesting and speech development of the picture exchange communication system in children with characteristics of autism. Journal of autism and developmental disorders 34, 395–409.

Garbacz, S.A., McIntyre, L.L., Santiago, R.T., 2016. Family involvement and parent–teacher relationships for students with autism spectrum disorders. School Psychology Quarterly 31, 478.

Ghafurian, M., Lakatos, G., Tao, Z., Dautenhahn, K., 2020. Design and evaluation of affective expressions of a zoomorphic robot, in: International Conference on Social Robotics, Springer. pp. 1–12.

Ghanouni, P., Jarus, T., Zwicker, J.G., Lucyshyn, J., Mow, K., Ledingham, A., 2019. Social stories for children with autism spectrum disorder: Validating the content of a virtual reality program. Journal of autism and developmental disorders 49, 660–668.

Gobrial, E., 2018. The lived experiences of mothers of children with the autism spectrum disorders in egypt. Social sciences 7, 133.

Goh, D.H., Ang, R.P., Tan, H.C., 2008. Strategies for designing effective psychotherapeutic gaming interventions for children and adolescents. Computers in Human Behavior 24, 2217–2235.

Gray, C.A., Garand, J.D., 1993. Social stories: Improving responses of students with autism with accurate social information. Focus on autistic behavior 8, 1–10.

Greczek, J., Kaszubski, E., Atrash, A., Matarić, M., 2014. Graded cueing feedback in robot-mediated imitation practice for children with autism spectrum disorders, in: The 23rd IEEE international symposium on robot and human interactive communication, IEEE. pp. 561–566.

Greenland, S., Senn, S.J., Rothman, K.J., Carlin, J.B., Poole, C., Goodman, S.N., Altman, D.G., 2016. Statistical tests, p values, confidence intervals, and power: a guide to misinterpretations. European journal of epidemiology 31, 337–350.

Gresham, F.M., Sugai, G., Horner, R.H., 2001. Interpreting outcomes of social skills training for students with high-incidence disabilities. Exceptional children 67, 331–344.

Groba, B., Canosa, N., Concheiro-Moscoso, P., 2018. Guidelines to support graphical user interface design for children with autism spectrum disorder: An interdisciplinary approach. Multidisciplinary Digital Publishing Institute Proceedings 2, 1177.

Gunning, C., Holloway, J., Fee, B., Breathnach, Ó., Bergin, C.M., Greene, I., Bheoláin, R.N., 2019. A systematic review of generalization and maintenance outcomes of social skills intervention for preschool children with autism spectrum disorder. Review Journal of Autism and Developmental Disorders , 1–28.

Habash, M., 2014. Emotion recognition and social skills in virtual environments for people with autism .

Hadwin, J., Baron-Cohen, S., Howlin, P., Hill, K., 1996. Can we teach children with autism to understand emotions, belief, or pretence? Development and psychopathology 8, 345–365.

HansonRobotics, 2005. Zeno research topic. Available at `https://www.hansonrobotics.com/zeno/?utm_source=robots.ieee.org`. [Accessed 01 January 2020].

Happé, F., Booth, R., Charlton, R., Hughes, C., 2006. Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: examining profiles across domains and ages. Brain and cognition 61, 25–39.

Herrero, J.F., Lorenzo, G., 2019. An immersive virtual reality educational intervention on people with autism spectrum disorders (asd) for the development of communication skills and problem solving. Education and Information Technologies , 1–34.

Howlin, P., Baron-Cohen, S., Hadwin, J., 1999. Teaching children with autism to mind-read: A practical guide for teachers and parents. J. Wiley & Sons Chichester.

Hughes-Roberts, T., Brown, D., 2015. Implementing a robot-based pedagogy in the classroom: initial results from stakeholder interviews, in: 2015 International Conference on Interactive Technologies and Games, IEEE. pp. 49–54.

Huijnen, C.A., Lexis, M.A., Jansens, R., de Witte, L.P., 2017. How to implement robots in interventions for children with autism? a co-creation study involving people with autism, parents and professionals. Journal of autism and developmental disorders 47, 3079–3096.

Huijnen, C.A., Verreussel-Willen, H.A., Lexis, M.A., de Witte, L.P., 2021. Robot kaspar as mediator in making contact with children with autism: A pilot study. International Journal of Social Robotics 13, 237–249.

Hung, L., Liu, C., Woldum, E., Au-Yeung, A., Berndt, A., Wallsworth, C., Horne, N., Gregorio, M., Mann, J., Chaudhury, H., 2019. The benefits of and barriers to using a social robot paro in care settings: a scoping review. BMC geriatrics 19, 1–10.

Ingersoll, B., 2008. The social role of imitation in autism: Implications for the treatment of imitation deficits. Infants & Young Children 21, 107–119.

Ingersoll, B., Lalonde, K., 2010. The impact of object and gesture imitation training on language use in children with autism spectrum disorder. Journal of Speech, Language, and Hearing Research .

Ingersoll, B., Schreibman, L., 2006. Teaching reciprocal imitation skills to young children with autism using a naturalistic behavioral approach: Effects on language, pretend play, and joint attention. Journal of autism and developmental disorders 36, 487.

Jarrold, W., Mundy, P., Gwaltney, M., Bailenson, J., Hatt, N., McIntyre, N., Kim, K., Solomon, M., Novotny, S., Swain, L., 2013. Social attention in a virtual public speaking task in higher functioning children with autism. Autism Research 6, 393–410.

Kaat, A.J., Lecavalier, L., 2014. Group-based social skills treatment: a methodological review. Research in Autism Spectrum Disorders 8, 15–24.

Kamaruzaman, M.F., Rani, N.M., Nor, H.M., Azahari, M.H.H., 2016. Developing user interface design application for children with autism. Procedia-Social and Behavioral Sciences 217, 887–894.

Kanner, L., et al., 1943. Autistic disturbances of affective contact. Nervous child 2, 217–250.

Ke, F., Moon, J., Sokolikj, Z., 2020. Virtual reality–based social skills training for children with autism spectrum disorder. Journal of Special Education Technology , 0162643420945603.

Khowaja, K., Banire, B., Al-Thani, D., Sqalli, M.T., Aqle, A., Shah, A., Salim, S.S., 2020. Augmented reality for learning of children and adolescents with autism spectrum disorder (asd): A systematic review. IEEE Access 8, 78779–78807.

King, A.M., Brady, K.W., Voreis, G., 2017. "it'sa blessing and a curse": Perspectives on tablet use in children with autism spectrum disorder. Autism & Developmental Language Impairments 2, 2396941516683183.

Kobak, K.A., Stone, W.L., Wallace, E., Warren, Z., Swanson, A., Robson, K., 2011. A web-based tutorial for parents of young children with autism: results from a pilot study. TELEMEDICINE and e-HEALTH 17, 804–808.

Kokina, A., Kern, L., 2010. Social story™ interventions for students with autism spectrum disorders: A meta-analysis. Journal of autism and developmental disorders 40, 812–826.

Koushki, R., Arabgol, F., Rahmani, M., Vahabie, A.H., et al., . Effectiveness of virtual reality-based therapeutic interventions on individuals with autism spectrum disorder: A comprehensive meta-analysis .

Kumazaki, H., Muramatsu, T., Yoshikawa, Y., Matsumoto, Y., Ishiguro, H., Kikuchi, M., Sumiyoshi, T., Mimura, M., 2020. Optimal robot for intervention for individuals with autism spectrum disorders. Psychiatry and Clinical Neurosciences 74, 581–586.

Lance, E.I., York, J.M., Lee, L.C., Zimmerman, A.W., 2014. Association between regression and self injury among children with autism. Research in Developmental Disabilities 35, 408–413.

Lawson, J., Baron-Cohen, S., Wheelwright, S., 2004. Empathising and systemising in adults with and without asperger syndrome. Journal of autism and developmental disorders 34, 301–310.

Lee, I.J., Lin, L.Y., Chen, C.H., Chung, C.H., 2018. How to create suitable augmented reality application to teach social skills for children with asd, in: State of the Art Virtual Reality and Augmented Reality Knowhow. IntechOpen.

Lee, J., Takehashi, H., Nagai, C., Obinata, G., Stefanov, D., 2012. Which robot features can stimulate better responses from children with autism in robot-assisted therapy? International Journal of Advanced Robotic Systems 9, 72.

Lindgren, S., Doobay, A., 2011. Evidence-based interventions for autism spectrum disorders. The University of Iowa, Iowa .

Lorenzo, G., Gómez-Puerta, M., Arráez-Vera, G., Lorenzo-Lledó, A., 2019. Preliminary study of augmented reality as an instrument for improvement of social skills in children with autism spectrum disorder. Education and Information Technologies 24, 181–204.

Lorenzo, G., Lledó, A., Pomares, J., Roig, R., 2016. Design and application of an immersive virtual reality system to enhance emotional skills for children with autism spectrum disorders. Computers & Education 98, 192–205.

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.

Lowry, L., 2011. Imitation with children on the autism spectrum: More than just a game of copycat.

Lux AI, 2019. Qtrobot-expressive robot helping children with autism learn emotions and positive social behaviors. Available at https://www.youtube.com/watch?v=68aesMHlHUs&t=5s. [Accessed 3 March 2020].

LuxAI, 2017. Qtrobot. Available at https://luxai.com/?utm_source=robots.ieee.org. [Accessed 01 January 2020].

LuxAI, 2021. Mental health benefits of teaching emotions to children with autism, using an expressive robot. https://luxai.com/blog/robot-enhanced-intervention-for-teaching-emotions-to-children-with-autism/. [Accessed 18 March 2021].

Maione, L., Mirenda, P., 2006. Effects of video modeling and video feedback on peer-directed social language skills of a child with autism. Journal of Positive Behavior Interventions 8, 106–118.

Makrygianni, M.K., Gena, A., Katoudi, S., Galanis, P., 2018. The effectiveness of applied behavior analytic interventions for children with autism spectrum disorder: A meta-analytic study. Research in Autism Spectrum Disorders 51, 18–31.

Malihi, M., Nguyen, J., Cardy, R.E., Eldon, S., Petta, C., Kushki, A., 2020. Data-driven discovery of predictors of virtual reality safety and sense of presence for children with autism spectrum disorder: A pilot study. Frontiers in psychiatry 11, 669.

Malinverni, L., Mora-Guiard, J., Padillo, V., Valero, L., Hervás, A., Pares, N., 2017. An inclusive design approach for developing video games for children with autism spectrum disorder. Computers in Human Behavior 71, 535–549.

Mandan, A., 2019. Autism classifications. https://www.news-medical.net/health/Autism-Classification.aspx5. [Accessed 14 July 2021].

Manfredonia, J., Bangerter, A., Manyakov, N.V., Ness, S., Lewin, D., Skalkin, A., Boice, M., Goodwin, M.S., Dawson, G., Hendren, R., et al., 2019. Automatic recognition of posed facial expression of emotion in individuals with autism spectrum disorder. Journal of autism and developmental disorders 49, 279–293.

Marino, F., Chilà, P., Sfrazzetto, S.T., Carrozza, C., Crimi, I., Failla, C., Busà, M., Bernava, G., Tartarisco, G., Vagni, D., et al., 2020. Outcomes of a robot-assisted social-emotional understanding intervention for young children with autism spectrum disorders. Journal of autism and developmental disorders 50, 1973–1987.

Marinoiu, E., Zanfir, M., Olaru, V., Sminchisescu, C., 2018. 3d human sensing, action and emotion recognition in robot assisted therapy of children with autism, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2158–2167.

Martin, F., Farnum, J., 2002. Animal-assisted therapy for children with pervasive developmental disorders. Western journal of nursing research 24, 657–670.

Masi, A., DeMayo, M.M., Glozier, N., Guastella, A.J., 2017. An overview of autism spectrum disorder, heterogeneity and treatment options. Neuroscience bulletin 33, 183–193.

Mechling, L., 2005. The effect of instructor-created video programs to teach students with disabilities: A literature review. Journal of Special Education Technology 20, 25–36.

MedCalc, 2021. User friendly statistical software. `https://www.medcalc.org/`. [Accessed April 2021].

Mesa-Gresa, P., Gil-Gómez, H., Lozano-Quilis, J.A., Gil-Gómez, J.A., 2018. Effectiveness of virtual reality for children and adolescents with autism spectrum disorder: an evidence-based systematic review. Sensors 18, 2486.

Miller, H.L., Bugnariu, N.L., 2016. Level of immersion in virtual environments impacts the ability to assess and teach social skills in autism spectrum disorder. Cyberpsychology, Behavior, and Social Networking 19, 246–256.

Miller-Wilson, K., 2006. Worksheets for children with autism. `https://autism.lovetoknow.com/Worksheets_for_Autistic_Children`. [Accessed 2 March 2020].

Miskam, M.A., Shamsuddin, S., Samat, M.R.A., Yussof, H., Ainudin, H.A., Omar, A.R., 2014. Humanoid robot nao as a teaching tool of emotion recognition for children with autism using the android app, in: 2014 International Symposium on Micro-NanoMechatronics and Human Science (MHS), IEEE. pp. 1–5.

Miskam, M.A., Shamsuddin, S., Yussof, H., Ariffin, I.M., Omar, A.R., 2015. A questionnaire-based survey: Therapist's response on emotions gestures using humanoid robot for autism, in: 2015 International Symposium on Micro-NanoMechatronics and Human Science (MHS), IEEE. pp. 1–7.

Mixamo, 2020. Get animated. Available at `https://www.mixamo.com/#/`. [Accessed 5 March 2020].

Moon, J., Ke, F., 2019. Exploring the treatment integrity of virtual reality-based social skills training for children with high-functioning autism. Interactive Learning Environments , 1–15.

Nojavanasghari, B., Hughes, C.E., Morency, L.P., 2017. Exceptionally social: Design of an avatar-mediated interactive system for promoting social skills in children with autism, in: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, ACM. pp. 1932–1939.

Ökcün-Akçamuş, M.Ç., Acarlar, F., Keçeli Kaysili, B., Alak, G., 2019. Examination of the relationship between gestures and vocabulary in children with autism spectrum disorder at different language stages. Early Child Development and Care 189, 777–791.

Ou, L.C., Luo, M.R., Woodcock, A., Wright, A., 2004. A study of colour emotion and colour preference. part i: Colour emotions for single colours. Color Research & Application 29, 232–240.

Özçalışkan, Ş., Adamson, L.B., Dimitrova, N., Baumann, S., 2017. Early gesture provides a helping hand to spoken vocabulary development for children with autism, down syndrome, and typical development. Journal of Cognition and Development 18, 325–337.

O'Handley, R.D., Radley, K.C., Whipple, H.M., 2015. The relative effects of social stories and video modeling toward increasing eye contact of adolescents with autism spectrum disorder. Research in Autism Spectrum Disorders 11, 101–111.

Panagiotidi, M., Wilson, S., Prescott, T., 2019. Exploring the potential of the animal-like robot miro as a therapeutic tool for children diagnosed with autism, in: Conference on Biomimetic and Biohybrid Systems, Springer. pp. 351–354.

Pavlov, N., 2014. User interface for people with autism spectrum disorders. Journal of Software Engineering and Applications 2014.

Peca, A., Tapus, A., Aly, A., Pop, C., Jisa, L., Pintea, S., Rusu, A., David, D., 2020. Exploratory study: Children's with autism awareness of being imitated by nao robot. arXiv preprint arXiv:2003.03528 .

Płatos, M., Wojaczek, K., 2018. Broadening the scope of peer-mediated intervention for individuals with autism spectrum disorders. Journal of Autism and Developmental Disorders 48, 747–750.

Poetker, 2019. What is virtual reality? (+3 types of vr experiences). `https://learn.g2.com/virtual-reality`. [Accessed 18 March 2019].

Pour, A.G., Taheri, A., Alemi, M., Meghdari, A., 2018. Human–robot facial expression reciprocal interaction platform: case studies on children with autism. International Journal of Social Robotics 10, 179–198.

Qi, C.H., Barton, E.E., Collier, M., Lin, Y.L., Montoya, C., 2018. A systematic review of effects of social stories interventions for individuals with autism spectrum disorder. Focus on Autism and Other Developmental Disabilities 33, 25–34.

Raisingchildren, 2018. Thinking and learning strengths in children with autism spectrum disorder. Available at https://raisingchildren.net.au/autism/learning-about-autism/about-autism/learning-strengths-asd. [Accessed 14 JUly 2021].

Raisingchildren, 2020. Social and emotional development. Available at https://raisingchildren.net.au/autism/development/social-emotional-development/emotional-development-asd. [Accessed 15 November 2020].

Raj, K.P., Sinha, S., Arvind, R.S., Solanki, D., Lahiri, U., 2018. Design of virtual reality based intelligent storytelling platform with human computer interaction, in: 2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS), IEEE. pp. 142–147.

Rao, P.A., Beidel, D.C., Murray, M.J., 2008. Social skills interventions for children with asperger's syndrome or high-functioning autism: A review and recommendations. Journal of autism and developmental disorders 38, 353–361.

Rasmussen, 2016. Flashback: Immersive virtual reality on mobile devices. https://medium.com/vr-today-magazine/flashback-immersive-virtual-reality-on-mobile-devices-40fab782eb45. [Accessed 18 July 2019].

Reyes, G.A., Tixi, V.E., Avila-Pesantez, D., Vaca-Cardenas, L., Avila, L.M., 2019. Towards an improvement of interpersonal relationships in children with autism using a serious game, in: The International Conference on Advances in Emerging Trends and Technologies, Springer. pp. 315–325.

Rice, L.M., Wall, C.A., Fogel, A., Shic, F., 2015. Computer-assisted face processing instruction improves emotion recognition, mentalizing, and social skills in students with asd. Journal of autism and developmental disorders 45, 2176–2186.

RobotLab2020, . Nao power v6 standard edition. Available at https://www.robotlab.com/store/nao-power-v6-standard-edition. [Accessed 18 September 2020].

Rogers, S.J., Wehner, E.A., Hagerman, R., 2001. The behavioral phenotype in fragile x: symptoms of autism in very young children with fragile x syndrome, idiopathic autism, and other developmental disorders. Journal of developmental & behavioral pediatrics 22, 409–417.

Rosenfield, N.S., Lamkin, K., Re, J., Day, K., Boyd, L., Linstead, E., 2019. A virtual reality system for practicing conversation skills for children with autism. Multimodal Technologies and Interaction 3, 1–9.

Rotheram-Fuller, E., Kim, M., Seiple, D., Locke, J., Greenwell, R., Levin, D., 2013. Social skills assessments for children with autism spectrum disorders. Autism 3, 1–8.

Rudovic, O., Lee, J., Dai, M., Schuller, B., Picard, R.W., 2018. Personalized machine learning for robot perception of affect and engagement in autism therapy. Science Robotics 3, eaao6760.

Rump, K.M., Giovannelli, J.L., Minshew, N.J., Strauss, M.S., 2009. The development of emotion recognition in individuals with autism. Child development 80, 1434–1447.

Salvador, M.J., Silver, S., Mahoor, M.H., 2015. An emotion recognition comparative study of autistic and typically-developing children using the zeno robot, in: 2015 IEEE International Conference on Robotics and Automation (ICRA), IEEE. pp. 6128–6133.

Sansosti, F.J., Powell-Smith, K.A., Kincaid, D., 2004. A research synthesis of social story interventions for children with autism spectrum disorders. Focus on autism and other developmental disabilities 19, 194–204.

Saracho, O.N., 2014. Theory of mind: children's understanding of mental states. Early Child Development and Care 184, 949–961.

Scattone, D., Tingstrom, D.H., Wilczynski, S.M., 2006. Increasing appropriate social interactions of children with autism spectrum disorders using social stories™. Focus on Autism and Other Developmental Disabilities 21, 211–222.

Schadenberg, B.R., Reidsma, D., Heylen, D.K., Evers, V., 2020. Differences in spontaneous interactions of autistic children in an interaction with an adult and humanoid robot. Frontiers in robotics and AI 7, 28.

Schuller, B., Marchi, E., Baron-Cohen, S., Lassalle, A., O'Reilly, H., Pigat, D., Robinson, P., Davies, I., Baltrusaitis, T., Mahmoud, M., et al., 2015. Recent developments and results of asc-inclusion: An integrated internet-based environment for social inclusion of children with autism spectrum conditions, in: Proceedings of the of the 3rd International Workshop on Intelligent Digital Games for Empowerment and Inclusion (IDGEI 2015) as part of the 20th ACM International Conference on Intelligent User Interfaces, IUI 2015, pp. 9–pages.

Sefidgar, Y.S., MacLean, K.E., Yohanan, S., Van der Loos, H.M., Croft, E.A., Garland, E.J., 2016. Design and evaluation of a touch-centered calming interaction with a social robot. IEEE Transactions on Affective Computing 7, 108–121.

Serret, S., Hun, S., Iakimova, G., Lozada, J., Anastassova, M., Santos, A., Vesperini, S., Askenazy, F., 2014. Facing the challenge of teaching emotions to individuals with low-and high-functioning autism using a new serious game: a pilot study. Molecular autism 5, 37.

Shahab, M., Taheri, A., Hosseini, S.R., Mokhtari, M., Meghdari, A., Alemi, M., Pouretemad, H., Shariati, A., Pour, A.G., 2017. Social virtual reality robot (v2r): a novel concept for education and rehabilitation of children with autism, in: 2017 5th RSI International Conference on Robotics and Mechatronics (ICRoM), IEEE. pp. 82–87.

Shen, J., Rudovic, O., Cheng, S., Pantic, M., 2015. Sentiment apprehension in human-robot interaction with nao, in: 2015 International Conference on Affective Computing and Intelligent Interaction (ACII), IEEE. pp. 867–872.

Shindorf, Z., 2016. Exploring barriers to the generalization of social skills interventions for children diagnosed with asd: A qualitative analysis of 'youth engagement through intervention' .

Shukla-Mehta, S., Miller, T., Callahan, K.J., 2010. Evaluating the effectiveness of video instruction on social and communication skills training for children with autism spectrum disorders: A review of the literature. Focus on autism and other developmental disabilities 25, 23–36.

Sigacheva, N., Baranova, A.R., Makaev, K.F., 2020. Communication problems: Advantages and disadvantages of teaching autistic children with humanoid robots. ARPHA Proceedings 3, 2387.

Smith, B.P., Dale, A.A., 2016. Integrating animals in the classroom: The attitudes and experiences of australian school teachers toward animal-assisted interventions for children with autism spectrum disorder. Pet Behaviour Science , 13–22.

So, W.C., Cheng, C.H., Lam, W.Y., Huang, Y., Ng, K.C., Tung, H.C., Wong, W., 2020. A robot-based play-drama intervention may improve the joint attention and functional play behaviors of chinese-speaking preschoolers with autism spectrum disorder: a pilot study. Journal of autism and developmental disorders 50, 467–481.

So, W.C., Wong, M.K.Y., Lam, C.K.Y., Lam, W.Y., Chui, A.T.F., Lee, T.L., Ng, H.M., Chan, C.H., Fok, D.C.W., 2018a. Using a social robot to teach gestural recognition and production in children with autism spectrum disorders. Disability and Rehabilitation: Assistive Technology 13, 527–539.

So, W.C., Wong, M.K.Y., Lam, W.Y., Cheng, C.H., Ku, S.Y., Lam, K.Y., Huang, Y., Wong, W.L., 2019. Who is a better teacher for children with autism? comparison of learning outcomes between robot-based and human-based interventions in gestural production and recognition. Research in developmental disabilities 86, 62–75.

So, W.C., Wong, M.K.Y., Lam, W.Y., Cheng, C.H., Yang, J.H., Huang, Y., Ng, P., Wong, W.L., Ho, C.L., Yeung, K.L., et al., 2018b. Robot-based intervention may reduce delay in the production of intransitive gestures in chinese-speaking preschoolers with autism spectrum disorder. Molecular autism 9, 1–16.

Soares, E.E., Bausback, K., Beard, C.L., Higinbotham, M., Bunge, E.L., Gengoux, G.W., 2021. Social skills training for autism spectrum disorder: A meta-analysis of in-person and technological interventions. Journal of Technology in Behavioral Science 6, 166–180.

Soares, F.O., Costa, S.C., Santos, C.P., Pereira, A.P.S., Hiolle, A.R., Silva, V., 2019. Socio-emotional development in high functioning children with autism spectrum disorders using a humanoid robot. Interaction Studies 20, 205–233.

Society, A., 2016. Facts and statistics. https://www.autism-society.org/what-is/facts-and-statistics/. [Accessed 26 July 2021].

Softbankrobotics, 2017. Softbankrobotics. Available at https://www.softbankrobotics.com/us/NAO. [Accessed 01 January 2020].

Sony, 2006. Aibo. Available at https://us.aibo.com/. [Accessed 11 July 2020].

Stanton, C.M., Kahn, P.H., Severson, R.L., Ruckert, J.H., Gill, B.T., 2008. Robotic animals might aid in the social development of children with autism, in: 2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI), IEEE. pp. 271–278.

Statsdirect, 2021. P values. https://www.statsdirect.co.uk/help/basics/p_values.htm. [Accessed April 2021].

Stewart Rosenfield, N., Lamkin, K., Re, J., Day, K., Boyd, L., Linstead, E., 2019. A virtual reality system for practicing conversation skills for children with autism. Multimodal Technologies and Interaction 3, 28.

Stone, W.L., Ousley, O.Y., Littleford, C.D., 1997. Motor imitation in young children with autism: What's the object? Journal of abnormal child psychology 25, 475–485.

Sutton, T.M., Altarriba, J., 2016. Color associations to emotion and emotion-laden words: A collection of norms for stimulus construction and selection. Behavior research methods 48, 686–728.

Taheri, A., Alemi, M., Meghdari, A., Pouretemad, H., Basiri, N.M., Poorgoldooz, P., 2015. Impact of humanoid social robots on treatment of a pair of iranian autistic twins, in: International Conference on Social Robotics, Springer. pp. 623–632.

The National Autistic Society, 2016. Autism. Available at https://www.autism.org.uk/about/what-is/asd.aspx. [Accessed 13 August 2020].

Thill, S., Pop, C.A., Belpaeme, T., Ziemke, T., Vanderborght, B., 2012. Robot-assisted therapy for autism spectrum disorders with (partially) autonomous control: Challenges and outlook. Paladyn 3, 209–217.

Trimmer, E., McDonald, S., Rushby, J.A., 2017. Not knowing what i feel: Emotional empathy in autism spectrum disorders. Autism 21, 450–457.

Tsikinas, S., Xinogalos, S., 2020. Towards a serious games design framework for people with intellectual disability or autism spectrum disorder. Education and Information Technologies , 1–19.

Ugwuanyi, L., Ubah, J., Obatta, M., Adama, G., Adaka, T., Nwabueze, J., Ejide, A., Otu, M.S., 2018. Using story telling technique in developing language in children with autism in nigeria: Perceptions and challenges. International Journal of Applied Engineering Research 13, 15343–15350.

UIUXTrend, 2021. Measuring and interpreting system usability scale (sus). Available at https://uiuxtrend.com/measuring-system-usability-scale-sus/. [Accessed 24 May 2021].

UnityAssetStore, 2020. Jammo character. Available at https://assetstore.unity.com/packages/3d/characters/jammo-character-mix-and-jam-158456. [Accessed 1 January 2020].

UniversityofHertfordshire, 2005. Kaspar the social robot. Available at https://www.herts.ac.uk/kaspar/the-social-robot. [Accessed 01 January 2020].

Valadão, C.T., Goulart, C., Rivera, H., Caldeira, E., Bastos Filho, T.F., Frizera-Neto, A., Carelli, R., 2016. Analysis of the use of a robot to improve social skills in children with autism spectrum disorder. Research on Biomedical Engineering 32, 161–175.

Valencia, K., Rusu, C., Quiñones, D., Jamet, E., 2019. The impact of technology on people with autism spectrum disorder: a systematic literature review. Sensors 19, 4485.

Vasquez, E., Nagendran, A., Welch, G.F., Marino, M.T., Hughes, D.E., Koch, A., Delisio, L., 2015. Virtual learning environments for students with disabilities: A review and analysis of the empirical literature and two case studies. Rural Special Education Quarterly 34, 26–32.

Wideo, 2020. Wideo free text to speech software(tts). Available at https://wideo.co/text-to-speech/. [Accessed 03 June 2020].

Wijker, C., Spek, A., Leontjevas, R., Verheggen, T., Enders-Slegers, M., 2017. The effectiveness of animal assisted therapy in adults with autism spectrum disorder: Study protocol for a randomized controlled trial. Autism Open Access 7.

Williams, B.T., Gray, K.M., 2013. The relationship between emotion recognition ability and social skills in young children with autism. Autism 17, 762–768.

Wood, L.J., Robins, B., Lakatos, G., Syrdal, D.S., Zaraki, A., Dautenhahn, K., 2019. Developing a protocol and experimental setup for using a humanoid robot to assist children with autism to develop visual perspective taking skills. Paladyn, Journal of Behavioral Robotics 10, 167–179.

Xin, J.F., Sutman, F.X., 2011. Using the smart board in teaching social stories to students with autism. Teaching exceptional children 43, 18–24.

Zhang, L., Fu, Q., Swanson, A., Weitlauf, A., Warren, Z., Sarkar, N., 2018. Design and evaluation of a collaborative virtual environment (comove) for autism spectrum disorder intervention. ACM Transactions on Accessible Computing (TACCESS) 11, 1–22.

Zheng, Z., Bekele, E., Swanson, A., Weitlauf, A., Warren, Z., Sarkar, N., 2017. The impact of robots on children with autism spectrum disorder. Autism Imaging and Devices , 397.

Zorcec, T., Robins, B., Dautenhahn, K., 2018. Getting engaged: assisted play with a humanoid robot kaspar for children with severe autism, in: International Conference on Telecommunications, Springer. pp. 198–207.

Appendices

A Ethical Approval

UNIVERSITY of GREENWICH

Maha Hatem Mohamed Abdelmohsen
Faculty of Liberal Arts & Sciences
Greenwich Campus

Direct Line 020 8331 8860
Email researchethics@gre.ac.uk
Our Ref UREC/19.5.5.10
Date: 24 August 2020

Dear Maha,

University Research Ethics Committee – Application 19.5.5.10

TITLE OF RESEARCH:	Enhancing Social Skills of Autistic Children by Assistive Teaching

I am writing to confirm that the above application has been **approved** by Chair's Action on behalf of the Committee and that you have permission to proceed.

I am advised by the Committee to remind you of the following points:

- You must notify the Committee immediately of any information received by you, or of which you become aware, which would cast doubt upon, or alter, any information contained in the original application, or a later amendment, submitted to the Committee and/or which would raise questions about the safety and/or continued conduct of the research;
- You must comply with Data Protection legislation;
- You must refer proposed amendments to the protocol to the Committee for further review and obtain the Committee's approval thereto prior to implementation (except only in cases of emergency when the welfare of the subject is paramount).
- You are authorised to present this University of Greenwich Research Ethics Committee letter of approval to outside bodies in support of any application for further research clearance.

On behalf of the Committee may I wish you success in your project.

Yours sincerely

Peter Garrod
Secretary, University Research Ethics Committee

University of Greenwich
Greenwich Campus
Old Royal Naval College
Park Row
London SE10 9LS
Telephone: +44 (0)20 8331 8000

B Egyptian Ethical Approval

إدارة الأمن

السيد الأستاذ / مدير عام إدارة ...الهرم........ التعليمية

تحية طيبة وبعد ،،،

بناء على الطلب المقدم من الباحثة / مها حاتم محمد – المسجلة للحصول على الماجستير كلية الآداب - جامعة جرينتش - انجلترا - وموافقة الإدارة المركزية للأمن برقم ٢٠٩٨/١ بتاريخ ٢٠٢٠/١١/٩

بشأن الموافقة لها على إجراء دراسة ميدانية بعنوان :-

(التقليد والتعرف على المشاعر والإيماءات اللاإنسانية لدى الأطفال المصابين باضطراب طيف التوحد)

لذا نفيد علم سيادتكم بأن المديرية قد وافقت لها وانه ليس هناك ما يمنع من وجهة نظر الأمن مع مراعاة الاتى:-

١ - التأكد من شخصية الباحثة . ٢ - استيفاء البيانات الشخصية .

٣ - موافقة أولياء أمور مفردات العينة وعدم جمع اى بيانات شخصية .

٤- بما لا يؤثر على سير العملية التعليمية .

٥- يتم ذلك تحت إشراف مدير عام الإدارة ومسئول امن الإدارة و تحت إشراف التوجيه المختص ومدير المدرسة والأمن بها وطبقاً للقواعد المنظمة فى هذا الشأن .

وتفضلوا بقبول وافر التحية والاحترام ،،

مدير إدارة الأمن

علاء الدين صلاح الدين

يعتمد ،،،

مدير المديرية

C Pre-Questionnaire

Social Skills Questionnaire

Child's Code:________________________ **Age: ______**

Parent/teacher/Other (please specify):_______________________ **Date:_________________**

Please use the following scale to indicate how well your child does each of the following:

1	2	3	4
Never	Sometimes	Often	Almost/Always

	Item	1	2	3	4	
1.	Responds to questions					**Responding to Interaction**
2.	Follow instructions					
3.	Pay attention for the instructions					
4.	Is aware of others' thoughts and feelings					**Affective Understanding/ Perspective taking**
5.	Is able to correctly identify other's feelings from their facial expressions, voice tone and body postures					
6.	Is aware of his/her own thoughts and feelings					
7.	Understands other's basic emotions. Ex: happy, sad, etc					
8.	Understands the use of basic emotions					
9.	Uses gestures to express emotions					
10.	Uses conventional facial expressions to express his/her feelings. Ex: raised eyebrows to express surprise					
11.	Understands emotions through social situations					
12.	Uses wide range of gestures to communicate. Ex: right/left palm waving (Hello)					**Motor/Play Skills**
13.	Understands other's gestures					
14.	Imitates other's actions/gestures					
15.	Uses appropriate gestures in social contexts					
16.	Understands gestures in social situations					

Thank you very much for your help

D Observation Sheet

Observation Sheet

Child's Code:_______________________ **Age: ______**

Parent/teacher/Other (please specify):______________________ **Date:________________**

Item	**Observation**
Skill	• **Emotion Skill** • **Intransitive Gesture**
Phase	• **Phase I** • **Phase II** • **Phase III**
Test	• **Pre-test** • **Training** • **Post-test**
Understanding the instructions of the scenario	• **Never** • **Sometimes** • **Always**
Paying attention to the instructions	• **Never** • **Sometimes** • **Always**
Interacting well with the tool	• **Never** • **Sometimes** • **Always**
Showing positive emotions during the session. Ex: happy, excited	• **Never** • **Sometimes** • **Always**
Showing negative emotions during the sessions. Ex: bored, angry	• **Never** • **Sometimes** • **Always**
Responding to the robot questions	• **Never** • **Sometimes** • **Always**
Asking you for help (How many time)	
Repeating words out of context	• **Never** • **Sometimes** • **Always**
Look away will you are talking to him/her	• **Never** • **Sometimes** • **Always**
Look away will the robot giving instructions	• **Never** • **Sometimes** • **Always**
Duration of the session	
Score (if contains)	
What are the emotions that he/she couldn't recognise?	

What are the gestures that he/she couldn't recognise?	
Any other comments/observation	

E Post-Questionnaire

Social Skills Post-Questionnaire

Child's Code:________________________ **Age: ______**

Parent/teacher/Other (please specify):_______________________ **Date:_________________**

Please use the following scale to indicate how well your child does each of the following:

1 Never	2 Sometimes	3 Often	4 Almost/Always

	Item	1	2	3	4	
1.	Responds to questions					**Responding to Interaction**
2.	Follow instructions					
3.	Pay attention for the instructions					
4.	Is aware of others' thoughts and feelings					**Affective Understanding/ Perspective taking**
5.	Is able to correctly identify other's feelings from their facial expressions, voice tone and body postures					
6.	Is aware of his/her own thoughts and feelings					
7.	Understands other's basic emotions. Ex: happy, sad, etc					
8.	Understands the use of basic emotions					
9.	Uses gestures to express emotions					
10.	Uses conventional facial expressions to express his/her feelings. Ex: raised eyebrows to express surprise					
11.	Understands emotions through social situations					
12.	Uses wide range of gestures to communicate. Ex: right/left palm waving (Hello)					**Motor/Play Skills**
13.	Understands other's gestures					
14.	Imitates other's actions/gestures					
15.	Uses appropriate gestures in social contexts					
16.	Understands gestures in social situations					
17.	The child's emotion recognition skill improved over the sessions					**Emotion Recognition Training Programme**
18.	The emotion recognition scenarios used were appropriate to train the child emotion ability					

19.	I am satisfied with the emotion recognition scenarios that my child practised					
20.	I would be interested in continuing using these emotion recognition scenarios to train my child to emotion ability					
21.	The child can generalise emotion recognition skill that has been learned in this study to daily life					
22.	The child's imitation skill improved over the sessions					**Intransitive Gesture Training Programme**
23.	The child's play skills improved over the sessions					
24.	The child starts using the taught gestures in daily life					
25.	The child's gesture recognition skill improved over the sessions					
26.	The intransitive gesture scenarios used was appropriate to teach the child gesture skill					
27.	I am satisfied with the gesture scenarios that my child practised					
28.	I would be interested in continuing using these gesture scenarios to teach my child to more gestures					

Thank you very much for your help

F System Usability Scale (SUS)

The System usability Scale

Child's Code:________________________ **Age: ______**

Parent/teacher/Other (please specify):______________________ **Date:________________**

Please use the following scale to evaluate the tool usability

1	2	3	4	5
Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**

	Item	1	2	3	4	5
1.	I think that I would like to use this tool frequently					
2.	I found the tool unnecessarily complex					
3.	I thought the tool was easy to use					
4.	I think that I would need the support of a technical person to be able to use this tool					
5.	I found the various functions in this tool were well integrated					
6.	I thought there was too much inconsistency in this tool					
7.	I would imagine that most people would learn to use this tool very quickly					
8.	I found the tool very awkward to use					
9.	I felt very confident using the tool					
10.	I needed to learn a lot of things before I could get going with this tool					

Thank you very much for your help

G Gesture Recognition Observation Sheet

Gesture Recognition Observation

Child's Code: **Age:**

Date:

The aim of this scenario is to examine whether the child is able to recognise the gestures produced by the robot.

Item	Yes	No
The child was able to recognise "Hello" gesture		
The child was able to recognise "Yes" gesture		
The child was able to recognise "Not allowed" gesture		
The child was able to recognise "Good Job" gesture		
The child was able to recognise "Where" gesture		
The child was able to recognise "Stop" gesture		
The child was able to recognise "Come" gesture		
The child was able to recognise "Look at this" gesture		
The child was able to recognise "Me" gesture		
The child was able to recognise "Awesome" gesture		
The child was able to recognise "Hungry" gesture		

H Participant Information Sheet

UNIVERSITY of GREENWICH

Participant Information Sheet

The main idea of this research is to investigate whether social virtual robot could be as effective as the physical robot in enhancing the social skills of autistic children and to develop a tool that is widely accessible and can be easily used by parents or teachers to practice social skills of autistic children and help them to live a better life. The target skills are imitation, emotion recognition skills and recognition and production od intransitive gestures.

The proposed tool is a combination between virtual environment and social robot. The tool employs a 3D robot that interacts with the child through a training programme guided by a parent or teacher.

The aim of hand gesture imitation training programme is to improve the imitation skill of the child as imitation of gestures and body movements predicts their language outcome and improve their play skills. The emotion recognition training programme divided into three phases: Learning emotions (Phase I), Recognising emotions from body postures (Phase II) and Recognising emotions from social context (Phase III). The aim of the intransitive gesture training programme is training child to recognise (Phase I) , imitate gestures (Phase II) and produce them in an appropriate context (Phase III).

Towards the evaluation process, the data will be gathered through questionnaires, observation and webcam. All the gathered data will be used for further analysis, however it will be completely anonymous and once the research has been completed the records will be destroyed. You as a parent/ teacher will be asked to fill out a questionnaire (Pre-Post sessions) to evaluate the effectiveness of the proposed work and to evaluate the generalisation of the acquired knowledge from the VR environment to the real contexts.

Your child will participate in these interactive sessions and interact with the virtual robot through the social scenarios. The study will be for approximately 3 months, but you have all the right to withdraw yourself and your child at any time during the study. All the gathered data will be stored in a secure folder in the university computing facilities and once the research has been completed, the records will be completely destroyed.

Signature of researcher	Date: 15/06/2020
This project is supervised by: Dr. Yasmine Arafa Prof. Cornelia Boldyreff Dr. Jixin Ma	
Researcher's contact details (including telephone number and e-mail address): Telephone number: " Due to the current circumstances and the closure of the University" Skype for business is the available alternative : ma4465r@gre.ac.uk Email: m.h.abdelmohsen@greenwich.ac.uk	

I Consent Form

UNIVERSITY of GREENWICH

PARTICIPANT CONSENT FORM

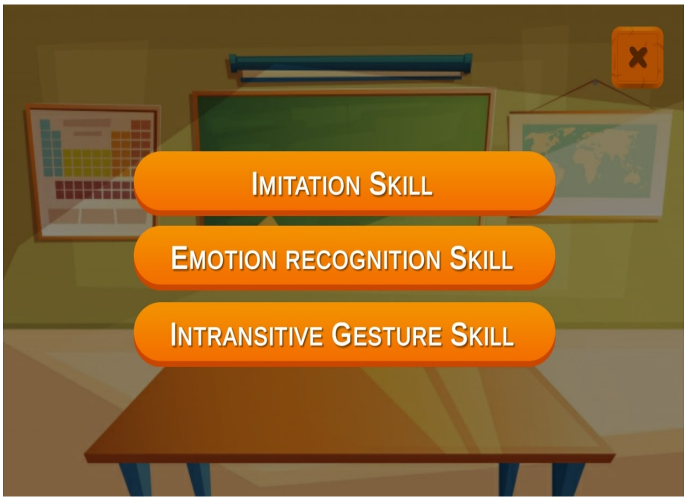
To be completed by the participant. If the participant is under 18, to be completed by the parent / guardian / person acting *in loco parentis*.

- I have read the information sheet about this study
- I have had an opportunity to ask questions and discuss this study
- I have received satisfactory answers to all my questions
- I have received enough information about this study
- I understand that I am / the participant is free to withdraw from this study:
  - At any time (until such date as this will no longer be possible, which I have been told)
  - Without giving a reason for withdrawing
  - (If I am / the participant is, or intends to become, a student at the University of Greenwich) without affecting my / the participant's future with the University
- I agree to take part in this study
- We may wish to use your research data for a further project in anonymous form. If you agree to this, please tick here ☐

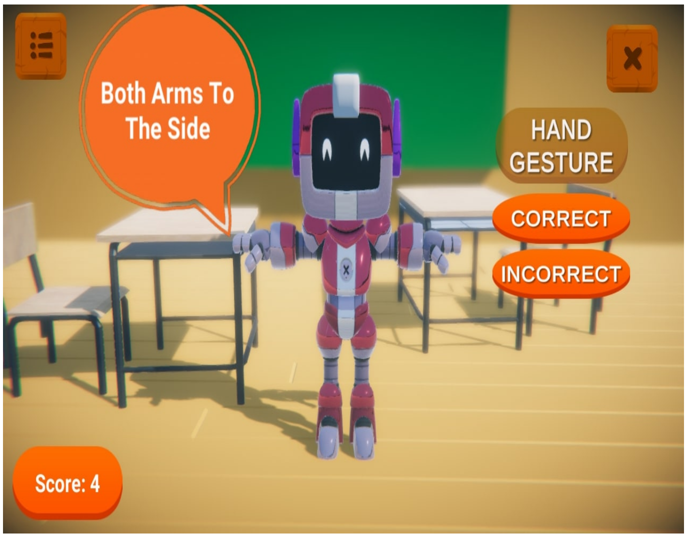
Signed (participant)	Date
Name in block letters	
Signed (parent / guardian / other) (if under 18)	Date
Name in block letters	
Signature of researcher	Date:
This project is supervised by: Dr. Yasmine Arafa Prof. Cornelia Boldyreff Dr. Jixin Ma	
Researcher's contact details (including telephone number and e-mail address): Telephone number: " Due to the current circumstances and the closure of the University" Skype for business is the available alternative : ma4465r@gre.ac.uk Email: m.h.abdelmohsen@greenwich.ac.uk	

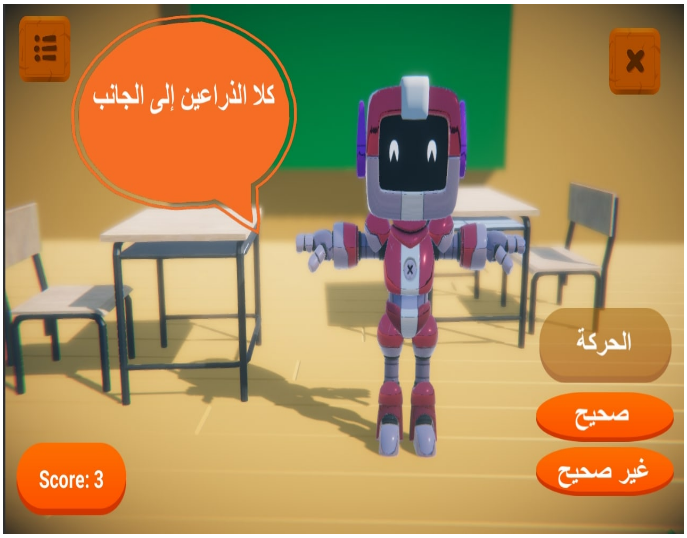
J Jammo-VRobot Environment

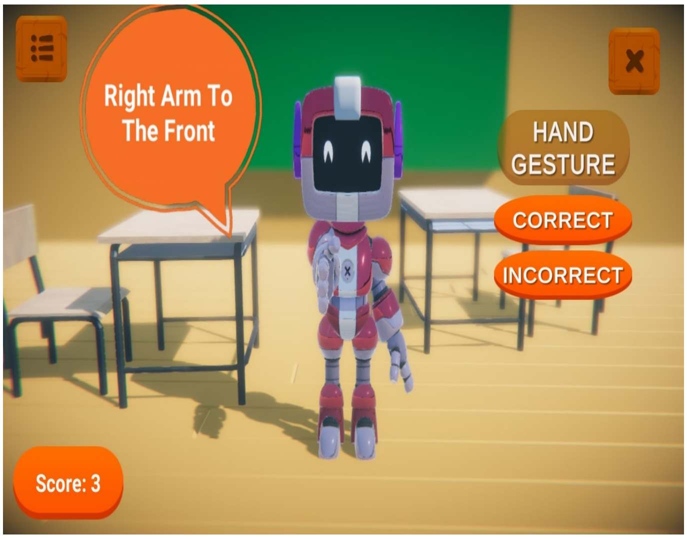
J.1 Menu



J.2 Imitation Scenario

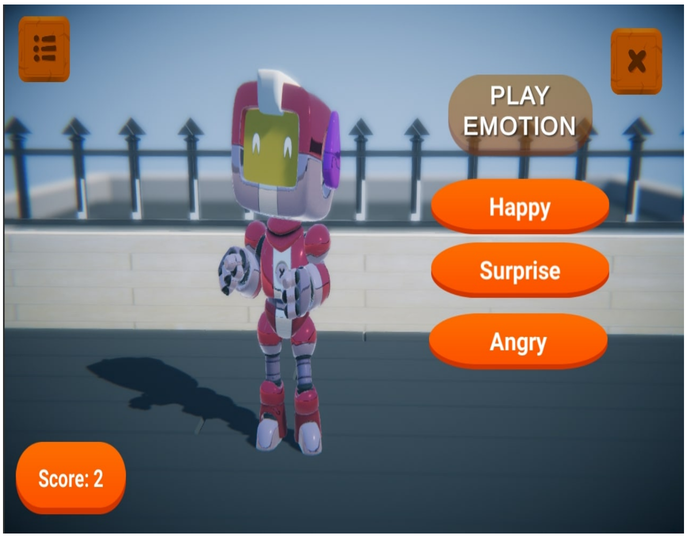


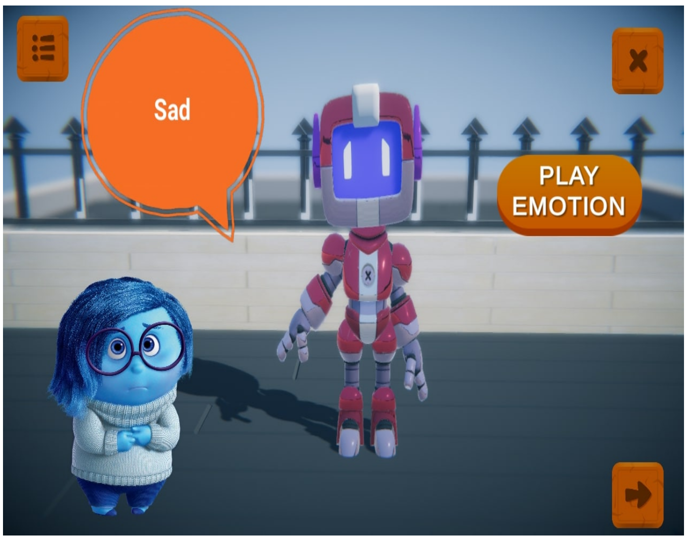


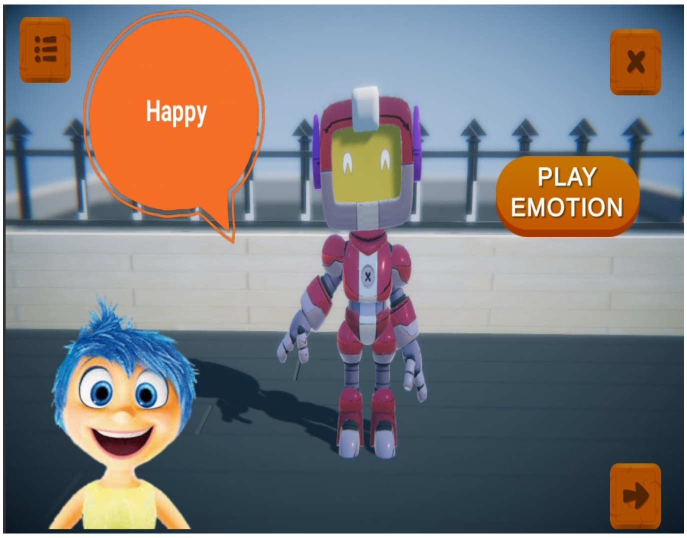


J.3 Emotion Recognition and Expression Scenarios

J.3.1 Phase I



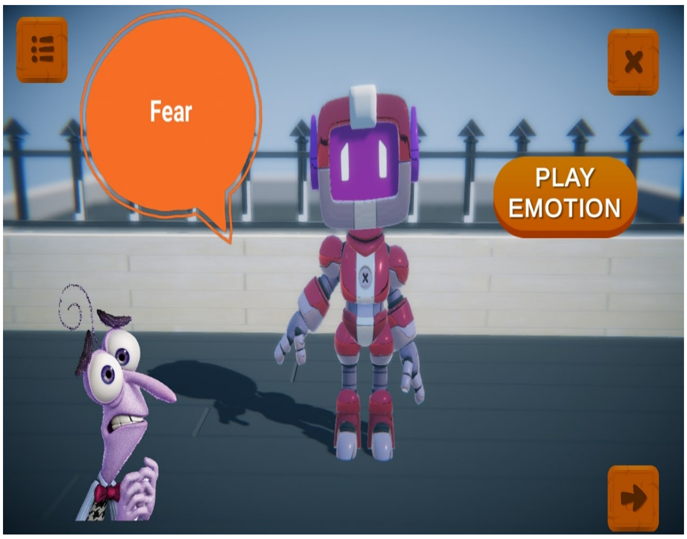


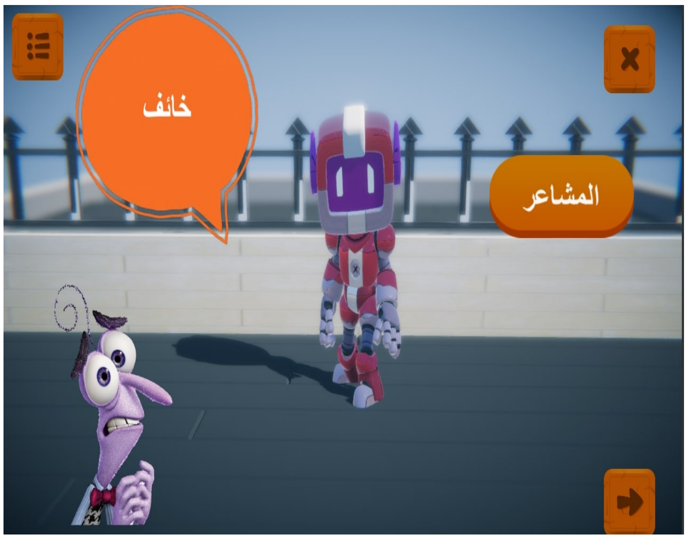

Happy
PLAY EMOTION


فرحان
المشاعر

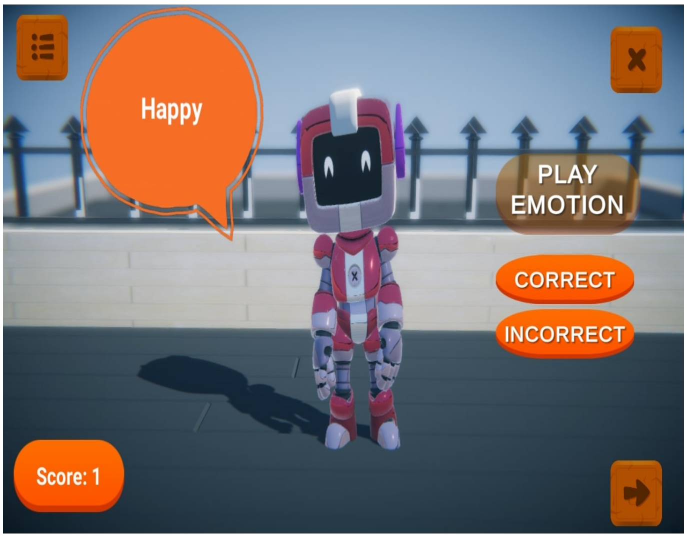

Angry
PLAY EMOTION

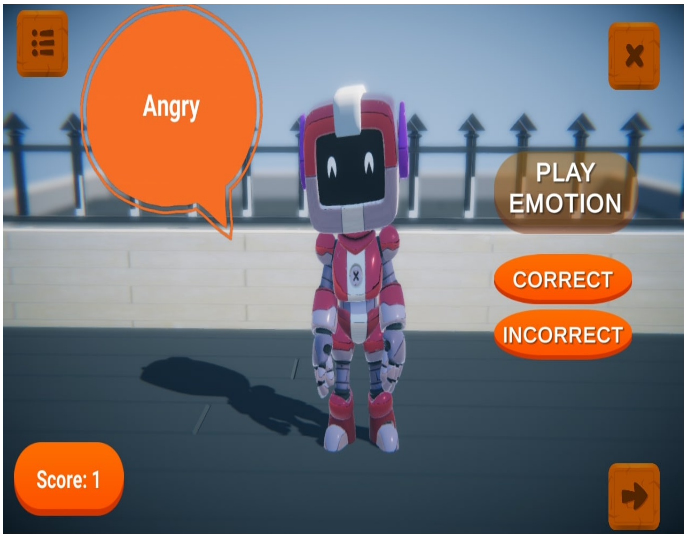

غضبان
المشاعر

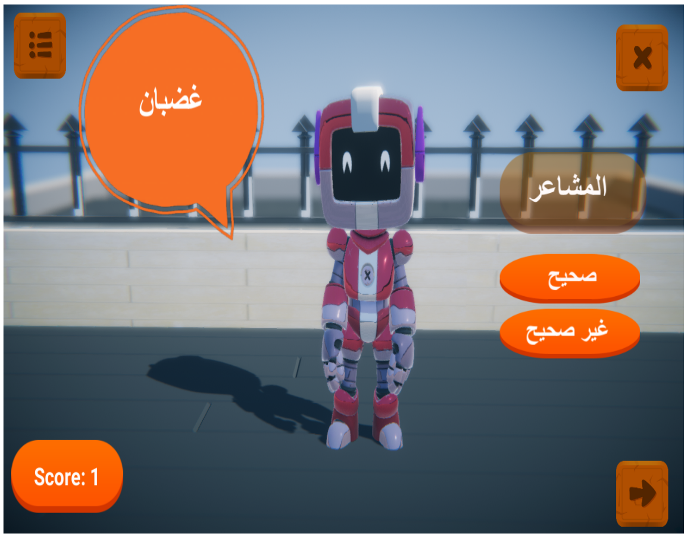

Fear
PLAY EMOTION


خائف
المشاعر

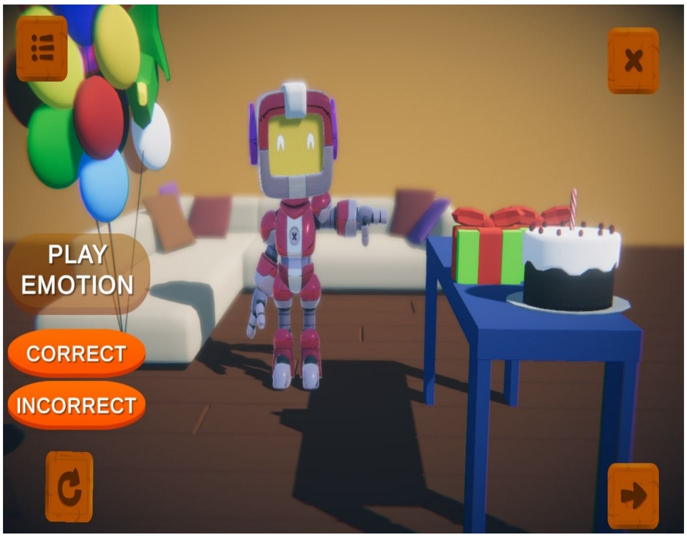
J.3.2 Phase II

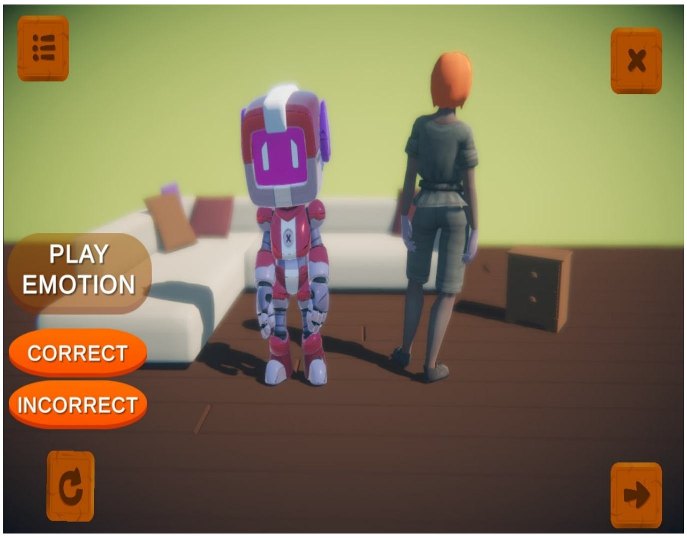






J.3.3 Phase III

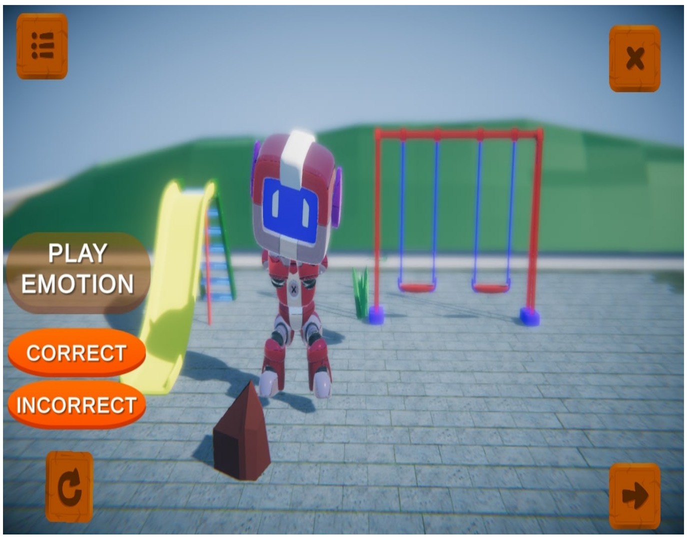



PLAY EMOTION
CORRECT
INCORRECT


المشاعر
صحيح
غير صحيح


PLAY EMOTION
CORRECT
INCORRECT

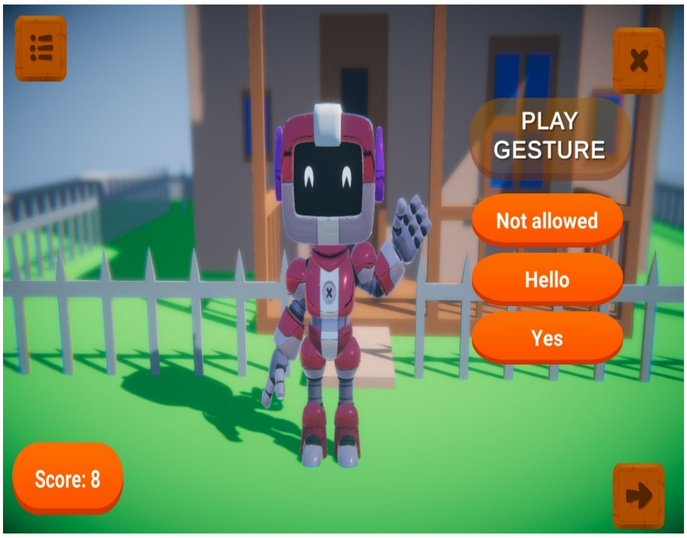

المشاعر
صحيح
غير صحيح

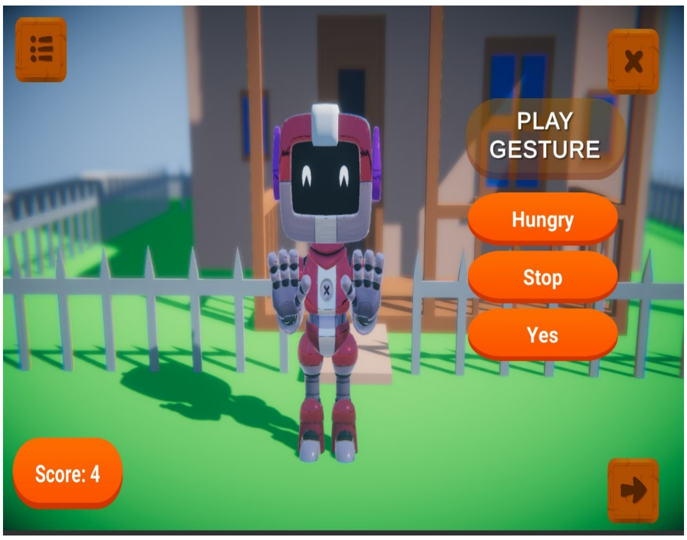

PLAY EMOTION
CORRECT
INCORRECT

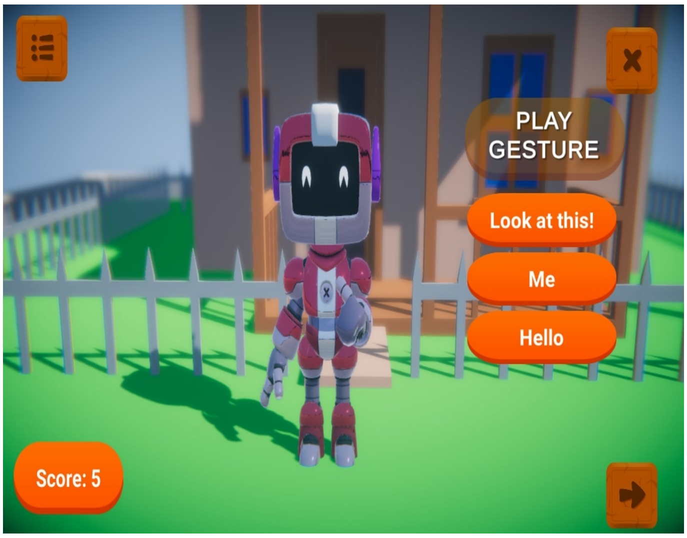

المشاعر
صحيح
غير صحيح

J.4 Intransitive Gesture Scenario

J.4.1 Phase I






PLAY GESTURE
Look at this!
Me
Hello
Score: 5


الإشارات
أنا
عمل جيد
غير مسموع
Score: 2


PLAY GESTURE
Look at this!
Stop
Good Job
Score: 1


الإشارات
مرحبا
أنا
عمل جيد
Score: 5


Not allowed
PLAY GESTURE


غير مسموع
الإشارة

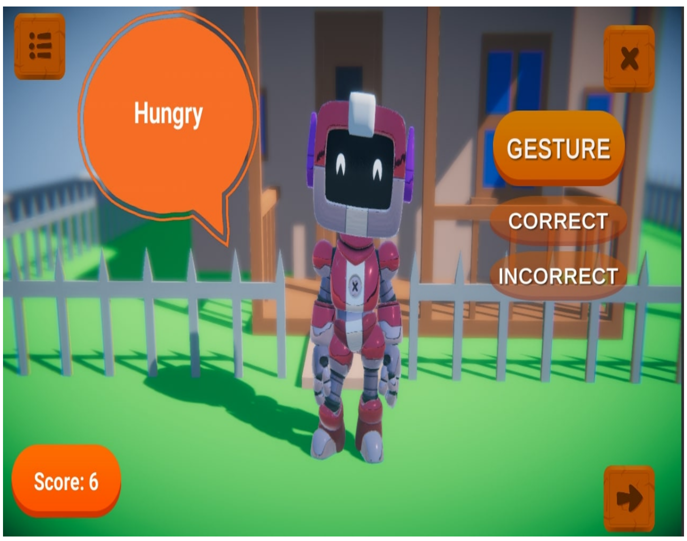

Stop
PLAY GESTURE


توقف
الإشارة

J.4.2 Phase II

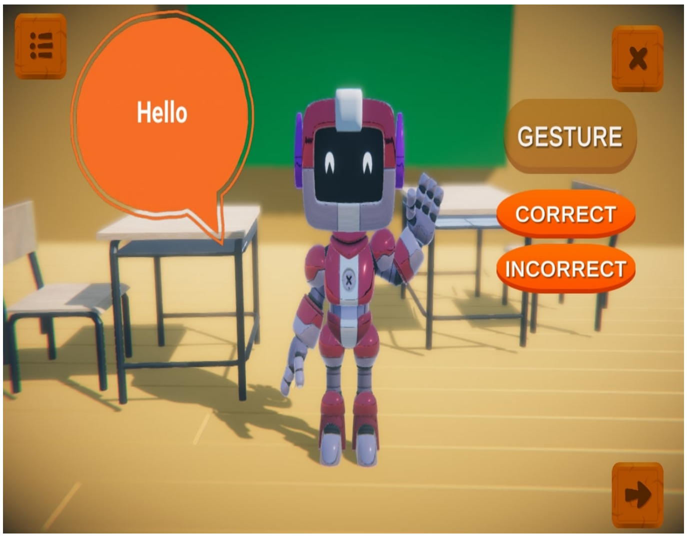

Me
GESTURE
CORRECT
INCORRECT
Score: 3

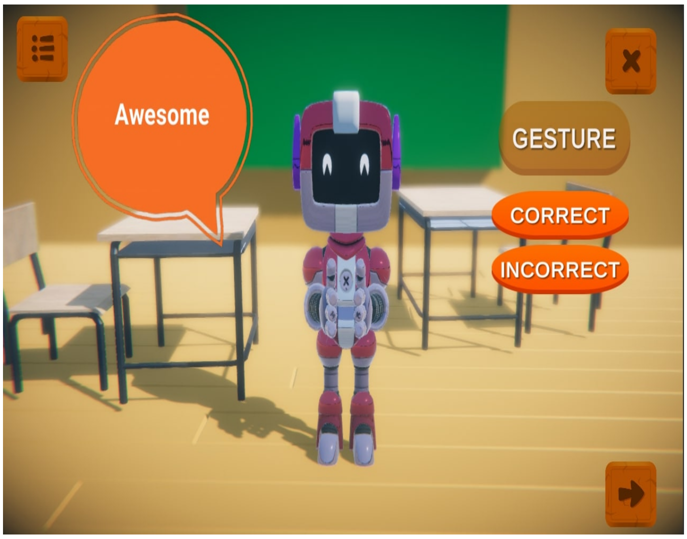

أنا
الإشارة
صحيح
غير صحيح
Score: 4

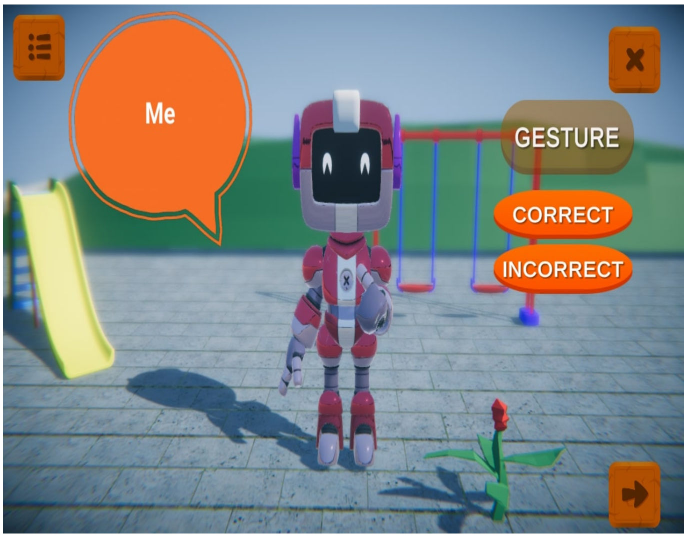

Hungry
GESTURE
CORRECT
INCORRECT
Score: 6

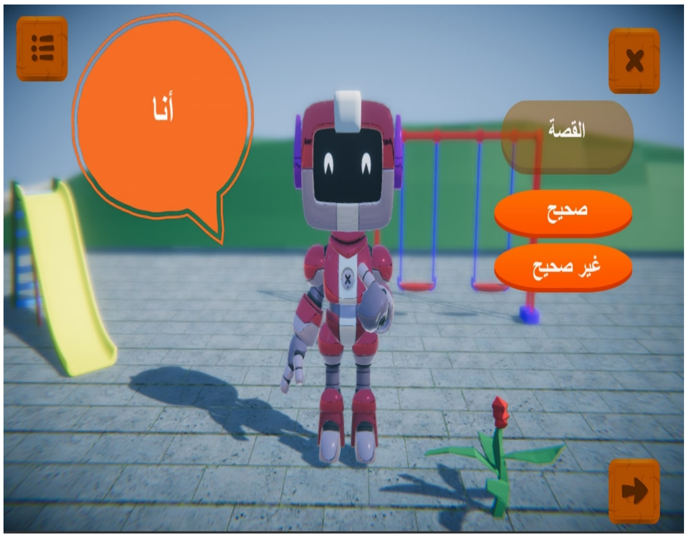

جعان
الإشارة
صحيح
غير صحيح
Score: 3

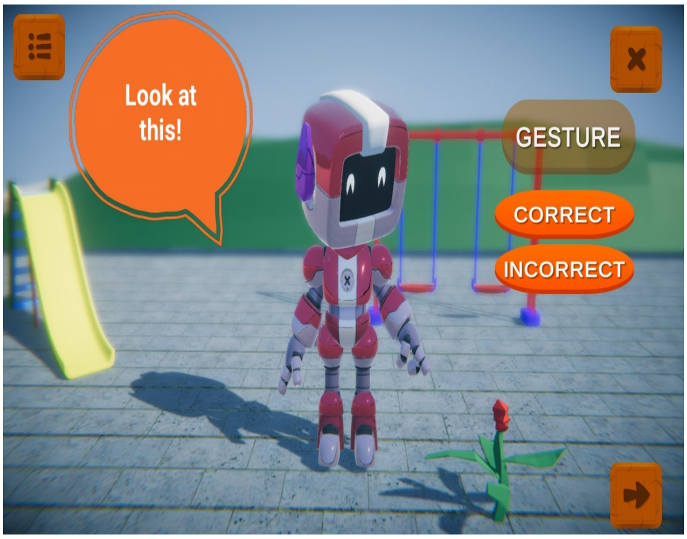
J.4.3 Phase III






Me
GESTURE
CORRECT
INCORRECT


أنا
القصة
صحيح
غير صحيح


Look at this!
GESTURE
CORRECT
INCORRECT


انظر الى هذا
القصة
صحيح
غير صحيح


Stop
GESTURE
CORRECT
INCORRECT


توقف
القصة
صحيح
غير صحيح