

Review

# Technology-Facilitated Diagnosis and Treatment of Individuals with Autism Spectrum Disorder: An Engineering Perspective

Xiongyi Liu <sup>1</sup>, Qing Wu <sup>2</sup>, Wenbing Zhao <sup>2,\*</sup> and Xiong Luo <sup>3</sup>

<sup>1</sup> Department of Curriculum and Foundations, Cleveland State University, Cleveland, OH 44115, USA; x.liu6@csuohio.edu

<sup>2</sup> Department of Electrical Engineering and Computer Science, Cleveland State University, Cleveland, OH 44115, USA; q.wu55@vikes.csuohio.edu

<sup>3</sup> School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China; xluo@ustb.edu.cn

\* Correspondence: w.zhao1@csuohio.edu; Tel.: +1-216-523-7480

Received: 8 July 2017; Accepted: 11 October 2017; Published: 13 October 2017

**Abstract:** The rapid development of computer and robotic technologies in the last decade is giving hope to perform earlier and more accurate diagnoses of the Autism Spectrum Disorder (ASD), and more effective, consistent, and cost-conscious treatment. Besides the reduced cost, the main benefit of using technology to facilitate treatment is that stimuli produced during each session of the treatment can be controlled, which not only guarantees consistency across different sessions, but also makes it possible to focus on a single phenomenon, which is difficult even for a trained professional to perform, and deliver the stimuli according to the treatment plan. In this article, we provide a comprehensive review of research on recent technology-facilitated diagnosis and treatment of children and adults with ASD. Different from existing reviews on this topic, which predominantly concern clinical issues, we focus on the engineering perspective of autism studies. All technology-facilitated systems used for autism studies can be modeled as human-machine interactive systems where one or more participants would constitute as the human component, and a computer-based or a robotic-based system would be the machine component. Based on this model, we organize our review with the following questions: (1) What are presented to the participants in the studies and how are the content and delivery methods enabled by technologies? (2) How are the reactions/inputs collected from the participants in response to the stimuli in the studies? (3) Are the experimental procedure and programs presented to participants dynamically adjustable based on the responses from the participants, and if so, how? and (4) How are the programs assessed?

**Keywords:** autism spectrum disorder; virtual reality; avatars; social robots; depth sensors; affective computing; emotion recognition; joint attention

---

## 1. Introduction

Autism Spectrum Disorder (ASD) is a neurological and developmental disorder. ASD presents with a constellation of physiologic and behavioral symptoms. There are two essential domains of ASD symptoms: (1) restrictive and repetitive behaviors, for example, having a lasting and intense interest in certain topics, including sensory challenges; (2) difficulty in social communication and interaction, such as failing to make proper eye contact during a conversation with another person. Although the exact causes for ASD are not known and there is no cure, research has shown that symptoms can be improved via proper treatment, especially when detected in early childhood.

It was estimated that 1.1% of population in the United States suffer from Autism Spectrum Disorders (ASD), which amounts to more than three million people. The cost of supporting

an individual with an ASD ranges from \$1.4 million for those without intellectual disability to \$2.3 million for those with intellectual disability during his/her lifespan (<https://www.autismspeaks.org/about-us/press-releases/annual-cost-of-autism-triples>). The largest cost components for children with ASD were special education services and parental productivity loss. Although the final cost is high for people with ASD, the actual damages caused by ASD are far beyond that. Both caregivers and people with ASD can be severely stressed due to the difficulty of proper mutual communication because people with ASD are often overwhelmed and anxious in social interactions. The lack of social skills for people with ASD will also significantly impact their personal lives and hamper their (future) career opportunities, especially for children. Many children with ASD have no friends and even have difficulty interacting with their parents. Not only do they often suffer from strong loneliness, but the frustration may also lead to violent behaviors, which will put both themselves and their loved ones in danger.

The rapid development of computer and robotic technologies in the last decade is giving hope to perform earlier and more accurate diagnosis of ASD, and more effective, consistent, and cost-conscious treatment. Besides the reduced cost, the main benefit of using technology to facilitate treatment is that stimuli produced during each session of the treatment can be controlled, which not only guarantees consistency across different sessions, but also makes it possible to focus on a single phenomenon, which is difficult even for a trained professional to perform, and deliver the stimuli according to the treatment plan. Computer based research typically uses virtual avatars or characters to represent therapists and other related facilitators, while robotic based treatment research typically uses humanoid robots. The computer based approach has the advantage of producing virtual reality scenes and scenarios for a treatment study, while the robotic based approach has the advantage of offering a real-human like physical embodiment.

In this article, we provide a comprehensive review of research on recent technology-facilitated diagnosis and treatment of children and adults with ASD. There are several recent reviews on similar topics. Wieckowski and White [1] provided an excellent review on technology-based interventions for children with social communication impairment. They reviewed work in the most recent seven years. They categorized the research into two major areas: (1) the reception of technology-based communication, and (2) the production of technology-based communication. For each area, the work is further divided into facial communication and non-facial communication. The research is predominately on facial communication. Non-facial communication includes verbal and gestures. The technologies that have been used include computer, mobile, virtual reality, and robotics. The discussions are focused on the clinical impact of these research works.

Aresti-Bartolome and Garcia-Zapirain reviewed the technology facilitated research work on ASD care [2]. They divided the technologies used into four categories: (1) virtual reality; (2) dedicated applications; (3) telehealth applications, and (4) robotic based. For each technology category, three research objectives are considered: (1) communication and interaction; (2) social learning and imitation skills; and (3) other miscellaneous.

There are several reviews on using a particular technology, such as virtual technology, social robotics, and serious games, to treat individuals with ASD. Scassellati et al. [3] reviewed about ten-year's work on using socially assistive robots for autism research up to 2012. The focus was on issues related to robot design, human-robot interaction, and system evaluation. The authors specifically highlighted the need to achieve robot autonomy because the operator-controlled method has severe limitations. Diehl et al. reviewed 15 papers that reported clinical use of social robots for individuals with ASD as of 2012 [4]. This review focused on four aspects of the studies, including the response of individuals with ASD to robots, the various ways of using robots to elicit behaviors, and using robots to teach social skills, and using robots to provide feedback. The authors concluded that most of the work reviewed were exploratory and had severe limitations. Even though it is not our focus, we find the situation has not been improved in our current review. Cabibihan et al. [5] outlined robot design features, compiled a list of robotic systems that have been used for ASD therapy up to 2013,

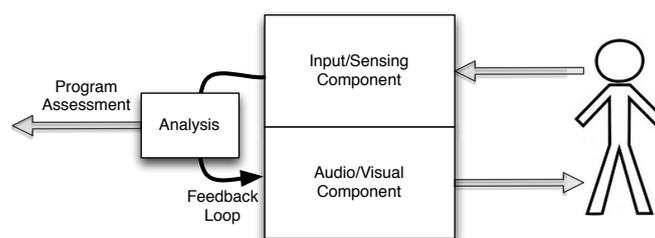
and highlighted a set of activities used for ASD therapy, including imitation, eye contact, joint attention, turn-taking, emotion recognition and expression, self-initiated interactions, and triadic interactions. The article also summarized the roles that could be played by a social robot, including as a diagnostic agent, as a friendly playmate, as a behavior eliciting agent, as a social mediator, as a social actor, and as a personal therapist. Pennisi et al. [6] provided a systematic review of research for autism treatment using social robots from a clinical perspective. A major contribution of this review is that it formulated ten questions regarding the essential roles and benefits of robots for ASD treatments, and reported their findings based on the studies that they have collected up to 2015 in light of these questions. Begum et al. [7] reviewed research on using social robots for autism treatment up to 2016. They pointed out that there is little progress in making social robots as clinical useful for ASD treatment.

Parsons provided a review on autism research facilitated by virtual reality technology [8]. This review focuses on the impact of veridicality of the virtual reality technology in autism treatment under two main themes: (1) using virtual environments for learning and intervention for individuals with ASD to better prepare them to handle real world interactions; and (2) using virtual environments to study how individuals with ASD would behave under predefined social scenarios.

Zakari, Ma, and Simmons [9] presented a short review of serious games designed for children with ASD to improve their communication skills and social behavior. These games are designed to be played via touch-enabled mobile devices or via traditional computers using a keyboard and a mouse. Very recently, Grossard et al. provided a more up-to-date review on using serious games to teach individuals with ASD to improve their social interactions and emotion skills [10]. They focused on the target skills of the serious games, and the design principles of the games. They also paid attention to whether or not the games have been validated in a clinical population.

Different from the above reviews, which predominantly concern clinical issues, we focus on the engineering perspective of autism studies. All technology facilitated systems used for autism studies can be considered as a form of human machine interactive systems where one or more participants would constitute as the human component, and a computer-based or a robotic-based system would be the machine component. As shown in Figure 1, the machine can be further divided into two subcomponents: (1) one subcomponent to deliver the output of the machine, typically in terms of visual and possibly audio content; (2) the other subcomponent takes input from the human, which could be directly from the human via a keyboard/mouse or a touch-screen, or indirectly via one or more wearable sensing devices, such as those for heart rate or skin conductance, or one or more remote tracking systems, such as those for eye-tracking or for gesture/body language tracking. Depending on the sophistication level of the human machine system, there may exist a feedback loop where the output of the system is dynamically altered in response to the input to the system. The feedback loop can be pre-programmed into the system, or controlled manually by a trained clinician (typically referred to as the Wizard of Oz scheme [3]). Furthermore, even without a realtime feedback loop, the effectiveness of a program can be evaluated via the data collected by the system objectively via technology, or subjectively from a clinical professional. In some clinical-oriented studies, pre-test and post-test are administered on the social behaviors of participants to assess the effectiveness of the treatment programs using relevant clinically-proven standardized tests.

Based on the model shown in Figure 1, we organize our review with the following questions: (1) What are presented to the participants in the studies and how are the content and delivery methods enabled by technologies? (2) How are the reactions/inputs collected from the participants in response to the stimuli in the studies? (3) Are the experimental procedure and programs presented to participants dynamically adjustable based on the responses from the participants, and if so, how? and (4) How are the programs assessed?



**Figure 1.** A system model for technology-facilitated systems used for autism research.

### *Research Contribution*

With the rapid development of new technologies, especially information technology and the unprecedented prevalence of autism spectrum disorder among the younger generation, researchers have shown tremendous interest in the field of technology-facilitated diagnosis and treatment of ASD. While it is an emerging field with promising opportunities, it is also one with challenges that are not typically faced by engineers, given that even experts in psychology and medical science are not sure about the causes and best treatment of the many different forms of ASD. Thus timely review of research in this field is crucial for researchers to navigate in this relatively new territory with good understanding of what we have learned from previous research efforts. Compared to previous reviews on similar topics, our review makes unique contribution in several aspects.

First, by utilizing a variety of literature search engines including Web of Science and Google Scholar, it breaks the boundaries among several disciplines that are highly involved in technology-facilitated diagnosis and treatment of ASD (psychology, medical science, education, and engineering) to provide a summary of the collective research efforts by researchers from each discipline.

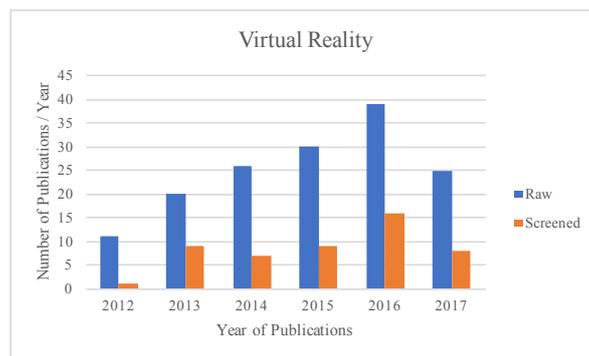
Second, by focusing on specific technology tools and applications, which serves as the common denominator, it overcomes the barriers caused by the idiosyncratic aspect of different terminology used in each discipline and presents a cross-disciplinary synthesis of the state of art research in the field that researchers can easily relate to regardless of their background. Unlike reviews that focus on specific domains, our review would facilitate more and more efficient sharing and co-construction of knowledge among researchers regarding technology-facilitated diagnosis and treatment of ASD across disciplines.

Third, our review takes into consideration that this is a relatively new and rapidly evolving field and aims at carving out the major areas of scientific exploration, identifying the areas that have been researched heavily as well as areas that demand more attention, evaluating existing evidence supporting the effectiveness of different technologies, and pointing out potentially fruitful directions that researchers can follow in the future.

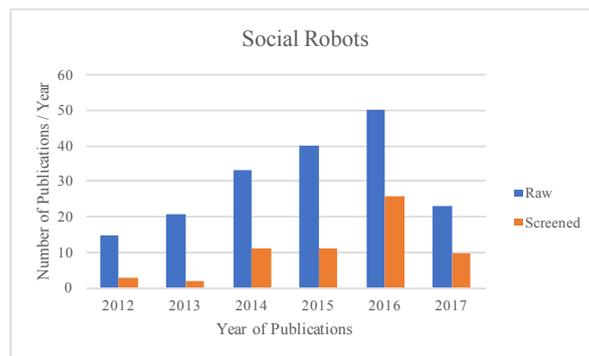
## **2. Inclusion of Literature**

We aim to provide a semi-systematic review of the studies that employ various technologies in the diagnosis and treatment of individuals with ASD. While it is desirable to have a rigorous systematic review of the topic, where all relevant publications are identified and carefully reviewed, several factors prevented us from doing so: (1) the review of numerous publications takes many months and during this period, new important publications were published and we do not think it is appropriate to ignore these latest publications; (2) there lacks a single search tool that can index all relevant publications and we do not have the manpower to exhaustively search for all databases for relevant publications. Therefore, we decided to rely on the use of Web of Science (all databases) as the main vehicle to find available publications and subjectively identify those that are relevant. We supplemented the search using Google Scholar to find relevant publications that are missing in Web of Science. We decided to focus on literature published since 2012 and only highly cited papers published earlier are included.

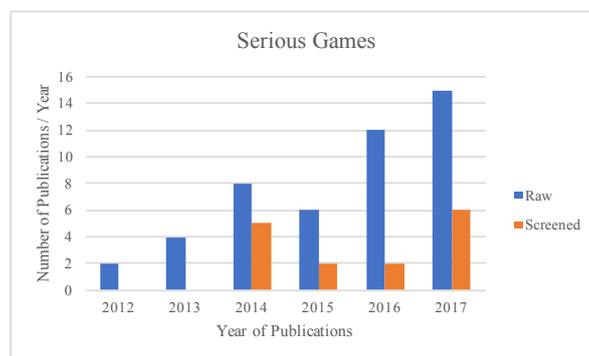
We primarily used three search terms, “serious games autism”, “virtual reality autism” and “social robots autism”. These terms help us obtain publications on autism treatment based on serious games, virtual reality technology, and social robots. The result for each search is manually screened to remove duplicates, misclassified papers (such as a paper on social robots is included in the search result for virtual reality) and out of scope papers (such as those not focused on autism). We selected papers on the diagnosis of individuals with ASD and those on helping individuals with ASD to improve their social skills and emotion regulation. We intentionally excluded papers that address other aspects of the lives of individuals with ASD, such as helping them to become more literate, to improve their locomotion skills, etc. In addition, we removed papers that do not contain sufficient details on the technology used in their studies. The search results (raw and screened) are summarized in Figure 2. It is interesting to note that the number of publications in each of the three areas are accelerating over the years.



(a)



(b)



(c)

**Figure 2.** Web of Science search results. (a) The raw and screened number of publications per year using the search term “virtual reality autism”; (b) The raw and screened number of publications per year using the search term “social robots autism”; (c) The raw and screened number of publications per year using the search term “serious games autism”.

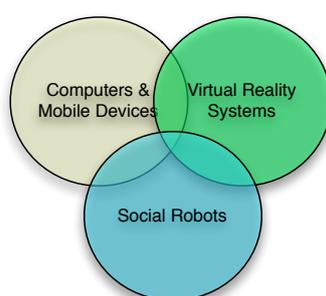
Based on the publications selected from the filtered Web of Science search results, we further searched Google Scholar to see if the same author(s) have any follow-up studies. If so, these follow-up studies are also included in this review.

The search results include both original research and review articles. We identified ten review papers, two on serious games for children with autism, one on virtual reality based treatment for autism, five for robotic based treatment for autism, and two comprehensive reviews that encompass all technology aspects.

### 3. Presentation to the Participants

All autism studies use one or more carefully crafted programs to diagnose, treat, or train a participating individual with ASD. The content of such a program is presented to a participant in certain ways, which could be as lightweight as a touch-enabled mobile device, a conventional computer monitor, an immersive virtual reality environment, or via a social robot. Each presentation method has its advantages and shortcomings in terms of cost, convenience, and effectiveness.

The above classification only considers the hardware forms used to deliver the content of a serious game or program to the participants. Except for social robots, the content delivered via all other means of content delivery is virtual in that it is digitally rendered. Even for social robots, some studies used ones that are equipped with a computer monitor for virtual head/face display instead of humanoid physical entity [11]. Furthermore, many studies that employed virtual reality programs used conventional computer monitors/TVs to delivery the content. Hence, the three major categories, i.e., computer and mobile devices, virtual reality systems/devices, and social robots, have certain degree of overlap, as shown in Figure 3.



**Figure 3.** Three major delivery forms of technology-facilitated autism studies.

After introducing the hardware based classification, we further consider user interface design issues in this section, in particular, the primary digital forms used to render the content for the serious games and programs, including images, videos, virtual avatars, non-human objects.

#### 3.1. Computers, Game Consoles and Mobile Devices

Most serious games designed for children with ASD are delivered via conventional computers (such as desktop or laptop computers) or mobile devices [12]. The visual content can be displayed on a laptop screen, a computer monitor, a TV screen, or a touch screen from a mobile device. An advantage of using a mobile device to deliver a game/program to a participant is the convenience [13–17]. A participant could play the game or engage in a program anytime he/she wants and virtually in any posture. Virtually all mobile devices and some newer computers (especially laptop computers) are equipped with touch screens, which enable users to make their choices directly on the screen instead of via a keyboard and mouse [18]. Some of the games designed to run on mobile devices [14] such as iPad, were adapted from traditional assistive tools used for ASD training, such as the cognitive affective training Kit (<http://www.cat-kit.com/>).

Some serious games involve avatars or a combination of avatars and pictures/videos of real people for various purposes such as the understanding emotions as well as recognition of different emotions via facial expression and body gestures [19]. Another interesting trend in serious games is development of collaborative virtual environment where users can collaborate and communicate during games [15,20,21].

In recent years, motion-based touchless serious games can be played using game consoles (selected from commercial games appropriate for children with ASD) [22–24], or via computer systems [25]. Different from traditional serious games, this type of games engage the full body movements of the participants. In addition to traditional learning targets such as social skills and attention skills, this type of games could enhance motor skills of children with ASD.

On the opposite end is a notion of tangible user interface where a user can directly touch physical objects, such as LEGOs. It is touted that the combination of tangible user interface and the computer-based graphical user interface will lead to more effective education/treatment for children with ASD [26].

Design guidelines for serious games for children with ASD were laid out in [25]. Even though it was introduced in the context of motion-based touchless games, they appear to be applicable for all serious games. In [27], Whyte et al. also provided a set of design principles for serious games. The guidelines include the following key points:

- Strong customizability. The game must be customizable to fit the needs and preferences of each child with autism. Unlike many diseases, autism children may have drastically different strengths and skill deficiencies.
- Increasing levels of complexity of game tasks. As a child acquires more skills, progressively more challenging tasks should be made available. This is the case for virtually all serious games. For example, Chua et al. [14] reported the development of an iPad-based game for children with ASD to learn emotion recognition. At the highest level, there are three worlds. There are six difficulty levels within each world. Each scenario contains one scenario depicted by a video of a human actor. At the end of the video, the headshot of the actor with a facial expression appropriate for the scenario is shown and the participant is asked to identify the emotion expressed by the headshot.
- Clear and easy to understand task goals. For each task, there should be a clear goal that can be easily understood by the participant.
- Multiple means of communicating game instructions, such as text, voice, and visual cues. Some children with low functioning ASD would need visual cues in particular.
- Positive reinforcement with rewards. Game score alone might not be enough to motivate children with ASD. Hence, other forms of rewards such as video or audio effect should be provided to encourage and motivate participants. In [13], a smiley face is shown at the completion of each game regardless of the score. This is referred to as reward-based intervention.
- Repeatability and predictability of game play. Unpredictability may cause anxiety to any children with ASD. The repeatability is needed for participants to learn.
- Smooth transitions. The game must be made easy to repeat and easy to transition to a higher level without noticeable delay so that children with ASD are not discouraged.
- Minimalistic graphics and sound/music. All graphical/sound elements must be included for the game goal because children with ASD may be subject to sensor overload. Even the use of color may play an important role for improved usability for children with ASD [14].
- Dynamic stimuli. Prolong static scene should be avoided to trigger motor rigidity.

### 3.2. Virtual Reality Systems/Devices

Three-dimensional immersive virtual reality can be realized via sophisticated immersive virtual reality system with surrounding displays, or head-mounted virtual reality devices. Some studies did use such virtual reality systems/devices [28–36]. However, many studies used conventional computer

monitors to display virtual reality scenes with one or more avatars [37–47]. Hence, there is certain degree of overlap between the virtual reality based presentation mode and the conventional computer based presentation mode. We should note that head-mounted display and virtual reality goggles may cause negative experiences such as dizziness and tiredness for some participants, as reported in [48]. Miller et al. provided an excellent review on the impacts of the level of immersion in virtual reality systems on ASD treatment [49].

In [28], a head-mounted display (eMagin Z800 3DVisor from eMagin Corporation at Bellevue, WA, USA) was used to present a virtual classroom containing multiple avatars to a participant. Different from virtual reality scenes provided by a conventional computer monitor, head-mounted display is capable of rendering the scenes in 3-dimension using stereoscopic images [50]. In addition to providing an immersive virtual reality environment, the head-mounted display can be integrated with head position monitor to detect eye gaze in the virtual visual field.

More recently, researchers started to experiment with using consumer-off-the-shelf virtual reality goggles for autism treatment. In [48], the Oculus Rift head-mounted display was used to study the willingness and acceptance of using the device, and the sense of presence and immersion the participants experienced with the device. Oculus Rift was also used in [51] to create immersive 3D scenes.

The term avatar has been used fairly liberally without a universal definition. In [25], Bartoli et al. divided the avatars used in their study into three categories: (1) articulated avatar, where the body parts are represented using lines and simple shapes, which is similar to a stick figure; (2) pointing avatar, where the body can move along a single direction; (3) real avatar, where the avatar is a silhouette of a real person. In other papers, a virtual avatar is typically referred to as a synthetic human character [52]. Carter et al. [52] further divided the avatar into two types: (1) animated avatar, which has an appearance of a 3-dimensional figure, (2) cartoon avatar, which is 2-dimensional like a cartoon character. Yet some literature calls a picture of a human face as an avatar [53]. Note that an avatar does not necessarily have to be human-like. In some programs, animal avatars were used, such as in the virtual pink dolphins project where avatars were dolphins [54–56].

While several different types of avatars, as well as pictures and videos of real persons have been used in many studies, few have compared their impacts to a participant regarding the effectiveness of the social skill learning. Carter et al. reported such a study [52]. In the study, two experiments were conducted. In both experiments, a therapist had a conversation with an individual with ASD. In the first experiment, three presentations of the therapist were used: (1) a realtime video feed of the therapist; (2) an animated avatar that tracks the facial motion of the therapist, and (3) a cartoon avatar that also tracks the facial motion of the therapist limited to 2-dimension. The result shows that there is no apparent variation among participants responses (verbal and nonverbal behaviors) under these three different modes of presentations. The second experiment focused on three different modes of the animated avatar. Exactly because the animated avatar is synthesized based on the therapist's facial expressions and motions, the avatar can be altered to the way the authors' desired. In this experiment, three conditions are tested: (1) direct facial motion tracking without any modification; (2) exaggerated facial motion; (3) dampened facial motion. The result shows that the participants exhibited more positive nonverbal responses for exaggerated facial motion, while there was no difference in verbal responses, which supports the authors' hypothesis that an animated avatar with more exaggerated facial expressions can enhance participant engagement and comprehension.

The virtual avatars used in most studies were produced using 3D animation platforms such as Unity or Vizard. Some studies chose to intentionally use 2D avatars either for comparison or to minimize sensory overload for participants with ASD. Yet some studies chose to use animated photos of real persons as avatars, such as FaceSay [53].

For virtual avatar based serious games and programs, it is essential to produce avatars that can be used to elicit stimuli appropriate for study. Undoubtedly, facial expressions are one of the most important elements in avatar production. In [52], the two-dimensional active appearance models

(AAMs) are used to track a therapist's facial motion. In the first step, AAMs are used to build a set of models based on the appearance and shape of the person's face. The learned models are used to produce an AAM mesh. This mesh is then used to track the facial motion of the person and to retarget to the avatar face.

In [28,37,39–41,57], the Vizard Virtual Reality Toolkit, a commercial product from WorldViz (<http://www.worldviz.com/vizard-virtual-reality-software/>) was used to produce the avatars in the study. The toolkit contains a PeopleMaker program, which users can use to custom design morphable 3D heads for avatars for different emotion expressions.

In [39], two different experimental paradigms, flashing paradigm and animated paradigm were used. The flashing paradigm involves three scenarios: (1) flashed schematic eyes, where the target event consists of eye balls slightly rotated while the non-target event consists of two eyeballs on the background screen; (2) flashed face with eye position change, where the target event consists of eyes gazing to the face's right and the non-target event consists of a face facing straight front; (3) flashed face with eye and head position change, where the target event consists of the face facing towards the left while the non-target event consists of a face facing direct forward. The animated paradigm involves with two scenarios: (1) animated 3D body with gaze change in one avatar, where the target event consists of the avatar head turning to the right side, and the non-target event consists of the avatar head turning to the left side; (2) animated avatars with gaze change in four avatars, where the scene contains four different avatars and the target event consists of the head of the avatar at the top row turning to the right while the non-target event consists of one of the avatars turning to the right.

In [40], the Vizard Virtual Reality Toolkit is also used. Instead of using the built-in avatars, they chose avatar heads from a database from the Virtual Human Interaction Lab at Stanford University. These avatar heads were converted from 2D photos of teenagers to 3D models using the 3DMeNow software (the company that created the software has ceased to exist). The authors further customized the complexion of the avatar heads using GIMP (GNU Image Manipulation Program), which is a free open source program. The study developed an emotion model that varies the expression from angry (the negative extreme) to happy (the positive extreme). In addition to face expression variations, an avatar can make gestures based on bone-kinematics, and walk around the scene.

In [28], the Vizard Virtual Reality Toolkit is used to develop a virtual 360-degree classroom that is delivered to a participant via a head-mounted display. Due to the 360-degree virtual environment, a participant would have to turn his/her head up to 60 degrees left or right to see all avatars in the scene. Vizard is also used in [44,58] to create virtual reality scenes.

In [59], the Game Studio A6 rendering engine from Conitec, Germany was used to produce the avatars and the corresponding scenes used in their study. In this study, the participants were asked to identify the emotion of the avatar, which could be happy, sad, surprise, anger, disgust, and fear. A participant was allowed to control the virtual distance between the participant and the avatar via joystick so that he/she could position the avatar to the most comfortable distance for emotion identification. For each emotion, there are four levels of emotional intensity (ranging from vague, neutral, to clear and strong) with corresponding facial expression (at 10, 40, 70, and 100%) and body gesture (such as "clasp hands, raise shoulder, shakes head, arms crossed, look down, step back" [59]). Furthermore, the rendering engine also enabled lip-synch animation to increase the fidelity of the verbal interaction.

Another platform used by researchers for virtual avatar-based scene production is the Unity game engine from Unity Technologies. Different from the Vizard Virtual Reality Toolkit and Game Studio A6, other modeling and animation tools are needed for dynamic facial expression design and animation. In [60], the Maya tool from Autodesk was used to animate the characters taken from Mixamo (<http://www.mixamo.com>). In this study, seven universal facial expressions were used, including "joy, surprise, contempt, sadness, fear, disgust, and anger" (page 1644) [60]. For each emotional expression, four emotion intensity levels were animated (low, medium, high, and extreme), which correspond to four different difficulty levels in facial expression recognition where the lower

intention level the more difficult to recognize. Unity is also used in [61] for virtual reality production, where both static and animated avatars are used and are capable of making ten different facial motions, including happiness, anger, disgust, fear, sadness, surprise, pain, neutral, and funny face. Unlike other studies, in [61], a game controller (referred to as gamepad in the paper) was used to control the avatars, control the cursors for answering questions, and receive vibration based tactile feedback.

In addition to facial expression, the gait could also reflect the emotion of a person. Hence, some studies have focused using polygonal walking avatars to produce emotional gait stimuli. Similar to facial expressions, gait can also be morphed into different intensity levels. In [62], three levels of intensity are used, ranging from attenuated (50%), prototypical (100%), and exaggerated (150%). The emotion can be expressed via the flexion of the head and arms, as well as the torso positioning [63]. For example, the happy emotion can be expressed by tilting the head backward and leaning back the torso, and the anger emotion can be expressed by tiling the head forward and leaning forward the torso.

In [64], the design of avatar has included the consideration that individuals with ASD tend to suffer from sensory overload, which refers to a phenomenon where an individual with ASD could choose to block all external stimuli when he/she is presented with too many cues concurrently. Hence, the authors of [64] chose to use a 2D cartoon-like avatar with predictable eye blink and head tilt patterns (the avatar links once every 5 s and tilts its head to the side once every 5 s as well).

In some studies, virtual avatars are placed in specific scenes that is consistent with the story told by the avatars. For example, if a story is about an incident in a fast food restaurant, the scene could be McDonald's. In [37,40,41], real 2D pictures of social environments were used for building the scenes. In other works, the scene is often designed using virtual 3D objects, such as [28], where a virtual classroom with multiple avatars is used in the study.

Because typically an avatar would tell a story verbally, the voice becomes an essential component in a virtual reality scene. The voices for avatars were recorded and replayed during a study session [37,41]. In some studies, lip sync, which can be facilitated by the Vizard Virtual Reality Toolkit, is used to make the story telling look more natural [40].

Yet some studies used the Second Life platform to create virtual reality scenes [46,65]. This approach could significantly reduce the development time for virtual reality scenes at the cost of flexibility.

### 3.3. Social Robots

Many studies have employed humanoid robots for treatment plans for children with ASD [66–79]. However, some research has used non-humanoid robots [11,80–84]. To employ humanoid robots in autism studies, one has to overcome two technical challenges: (1) how to control the movements of the robot, and (2) how to obtain the input from the participant. Unlike virtual avatars, where mature technologies exist to enable high fidelity animation of both facial expressions and body gestures, which are essential to facilitate closed-loop autonomous treatment programs, socially assistive robots used for autism studies so far are scripted in a very limited way (such as raising an arm etc.), presumably due to safety and cost concerns. Furthermore, the use of a social robot is meant to facilitate natural user interaction where the participant would engage in verbal conversations and/or use his/her gestures while doing so. In contrast, a virtual avatar based system can equip the participant with a keyboard and mouse, or a touch screen, which allows the participant to provide his/her input/response to a question by typing, clicking, and touching by hand.

There have been efforts towards an autonomous robotic system. Similar to a virtual reality based system, the eye gaze of the participant with respect to the robot can be detected using eye/head tracking devices, such as a remote desktop eye tracker [85] or a head-mounted eye tracking device [86]. Accurate tracking of eye gaze could enable partial autonomous operation of the treatment program via a social robot. Some studies have also enabled some form of speech recognition functionality. It is limited because for higher recognition accuracy, only scripted simple voice commands can be

recognized. There is still a long way to go to support unscripted conversations between a robot and a participant. There is also work on using cameras to track facial expressions [87,88] and gestures made by a participant to understand the emotion of the participant and to allow the participant to give commands to the robots for increased degree of interaction between robots and participants [11].

Despite recent progresses made, such as motion and eye tracking [89–91], in social robotic research, the Wizard of Oz scheme remains the dominant approach to give an illusion that the robot can capture and understand the participant input/response and react/adapt properly while in fact the expression and motion of the robot is controlled, and the user input is captured by a therapist and/or technician.

Many different types of robots have been employed in this line of research, including both humanoid robots and non-humanoid robots. Some robots are stationary in that they do not move around even though they can move their arms, legs and heads in certain ways. Some other robots are mobile in that they can be instructed to move around or even to follow the participant [11]. Most robots are equipped with cameras (for object recognition), microphones (for voice commands or speech recognition), and speakers. Some are equipped with tactile sensors at the head and head to sense touching.

Humanoid robot Nao from Aldebaran Robotics Company is one of the most popular robots used in autism studies [85,92–99]. Nao is a child-sized robot with plastic body (58 cm tall and 4.3 kg weight). It was used in [92] for joint attention and imitation study, in [85] for joint attention study, in [95] for interventions on repetitive behaviors and affective states, in [96] for interventions on verbal communication skills, and in [97,100] for intervention on imitation and interpersonal synchrony of children with ASD. In [95], an additional mobile robot called Rovio from WowWee was used.

Another humanoid robot Zeno R-30 was used in [101] to facilitate imitation training for children with ASD. Zeno R-30 is also a child-like robot produced by Hanson RoboKind. Zeno must be controlled by a separate computer. Zeno R-30 was used for an imitation study where the robot and the participant take turns imitating each other's movement facilitated by a Microsoft Kinect sensor, where the operator's movement was captured and mapped to the robot and similarly the participant's movement was captured for offline analysis.

A newer version of the Zeno robot called Robokind was used in [102,103] to engage children with ASD and emotion recognition training where the participant was asked to identify the emotion expressed by the robot. This robot has a walking body, which is a significant improvement over the Zeno robot. The robot's skin is made of a biomimetic polymer called Frubber, which enables it to make a rich facial expressions. Some research used both Nao and Zeno robots, such as [104].

Lucy is also a humanoid robot used in an autism study [105]. Lucy is about 39 centimeters tall and 6.5 kg in weight. It has a baby-like face and it is capable of talking, recognizing faces and tracking faces. It can make limited types of facial expressions by lighting up some LEDs in the face areas, and limited number of gestures. It is also equipped with motion sensors and touch sensors, and can perform some speech recognition. In [105], Lucy was used to play games (Bingo and quizzes) with children with ASD via a tablet. Lucy was programmed to say encouraging words when the participant provided corrected answers in the quizzes (via a touch panel connected to Lucy instead of speech recognition).

Non-humanoid robots have also been used for autism research. For example, Dehkordi et al. [106] developed a parrot-like robot called RoboParrot for screening autistic children. RoboParrot is placed on a fake wooden base inside a cage mounted with a webcam, with a USB cable coming out of its tail and connecting the embedded hardware with the micro-controller board. The hardware is partially embedded inside the body of RoboParrot and partially embedded in the base. Through this system, input data can be collected via the camera on the cage, and the microphone and various sensors embedded in the robot. Specifically, there are two sensors: an infrared sensor, which is in the chest of RoboParrot and detects if someone tries to touch the chest, and a Hall Effect sensor, which is placed in the beak of RoboParrot and senses the closeness of a person's hand over RoboParrot's head and break.

In [11], a non-traditional robot was designed for autism study. It is equipped with a computer monitor displaying a virtual avatar face and a depth sensor called Asus Xtion PRO for tracking the

gesture of the participant and for recognizing the face of the participant. This non-humanoid robot allows the participant to move it around via gestures, and engage in simple pre-scripted conversations via speed recognition and text to speech software. It is important to note that the usability of the robot was not validated via individuals with ASD.

It is interesting to note that social robots are sometimes used in conjunction with conventional computer systems. For example, in [85], two computer monitors are used (one on each side of a robot) in a program designed to improve the joint attention of individuals with ASD where objects were displayed on the monitors when the robot turns its head towards left or right. Furthermore, robotic studies are not just hardware-focused. There was also an effort to develop custom robots with multi-physiological sensing to automatically observe the emotion of a participant [107].

The need for interoperable software architecture for different brands of robots is also getting researchers' attention. In [108], efforts were made to design a platform-agnostic robot control architecture that can be used for various treatment scenarios.

#### 4. Input/Reactions from Participants

In response to stimuli introduced to a participant, various reactions from the participant are collected and often analyzed to different degrees either in realtime or offline. Some of these reactions are directly solicited from the participant where he/she are asked to provide explicitly manual input via a keyboard and mouse, a touch screen, or a specific action such as pointing to a particular object on the screen or touching a designated area in a social robot. Other studies use various sensing technologies, such as computer vision or physiological sensing, to capture more complete reactions produced by the participant, such as eye gaze, physiological characteristics, body language. Yet some studies depend on a trained professional to record reactions of the participant. These common means of getting the input and responses from participants are illustrated in Figure 4.

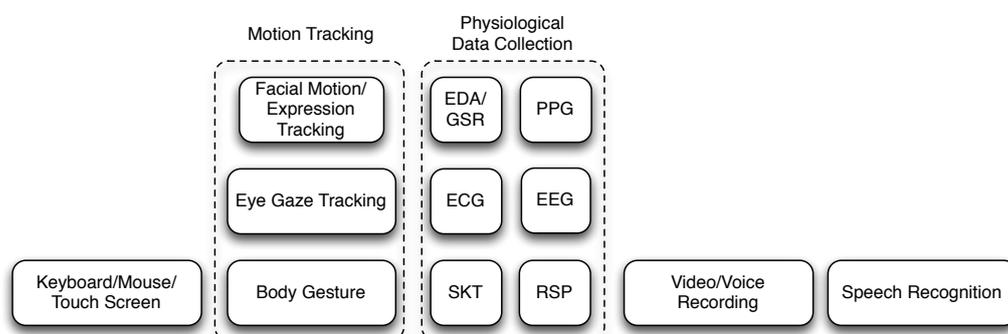


Figure 4. Common means of getting the input and responses from participants.

##### 4.1. Keyboard, Mouse, Touch Screen

To collect direct input from a participant, a conventional approach is to use a menu-driven interface or to display all possible options on the screen where a participant would choose the right answer from several response options using the mouse pointer [40,41,59–61]. It is also common to require a participant to type a key as the answer, for example, “A” for the anger emotion, and “H” for the happy emotion. It is interesting to note that not only the correctness of the selection or input is important, but in some research, the response time, i.e., the time it takes for a participant to make the selection/input since the prompt of the question is also recorded because this delay can be scientifically significant. For example, in [60], the response time was recorded to validate their hypothesis that it takes longer time for a child with ASD to recognize a facial affect than that for a typical developing child.

With the pervasiveness of mobile devices and laptops equipped with touch screen monitors, users could also provide direct input by touching a particular area on the touch screen. This is often the required input method for serious games delivered on devices with touch screens. In [53], participants are required to perform tasks via a touch screen in three different games, including selecting the objects that the avatar is looking at, selecting the right face band aid that would correct the avatar's face, select the correct facial expression, and change the expression of an avatar as suggested.

#### 4.2. Facial Motion/Expression Tracking

There are relatively mature computer vision technology to facilitate facial motion tracking, such as eye gaze, blinks, and mouth motions. For example, two-dimensional active appearance models [109] have been used for facial motion tracking in realtime. In some research [110], head tracking is implemented using a Kinect sensor as an approximation of eye gaze for joint attention study. We have also seen research on using facial expressions recognition to obtain affective feedback [30,111–116]. In some research, head orientation and movement are tracked using inertial sensors embedded in smart glasses instead of cameras [117]. While the tracking itself helps tremendously in producing realistic virtual avatar faces, there is still an open research issue in how to accurately understand the captured motion with respect to nonverbal communicative behaviors of a participant with the exception of eye gaze (and to some extent facial expressions). Hence, the state of the art practice is still to record the video of the participant and then use domain experts to analyze the video regarding the participant's nonverbal communicative behaviors [52,118–120].

#### 4.3. Eye Gaze Tracking

There are two major types of eye tracking technologies. The first one is head-mounted and the other is referred to as desktop based without requiring a participant to wear any device for eye tracking on his/her body. The core technology for eye tracking is identical to head-mounted and desktop eye trackers. An infrared camera is used to detect the rotations of the eyeballs (based on the pupil position and the corneal reflection). The reason why infrared light is used for eye movement tracking is that it renders a better demarcation of the pupil and the iris than visible light. Once the eye movements are determined, the gaze can be computed based on the context.

In [37,41,121], an eye-tracker goggles from Arrington Research Inc. was used to track eye gaze in realtime. The goggles are used in conjunction with a custom designed eye tracking interface platform. The raw eye gaze data were sampled 30 times a second. Then, three features were extracted from the raw data, including the mean pupil diameter, the mean blink rate, and the average fixation duration for each region of interest.

A head-mounted device called WearCam has been used to compute eye gaze offline. This device records the eyes of the participant and the image of the field of view in front of the participant simultaneously [122]. Based on the recorded data, the direction of gaze and focus of attention can be computed. A downside of WearCam is that it requires offline analysis of the recorded video to determine gaze. Hence, this device cannot be used if eye gaze needs to be determined in realtime.

Eye gaze tracking can be made simpler if used in conjunction with head-mounted virtual reality devices, which produces a 360-degree immersive virtual 3D scene. In this case, it is sufficient to track the head movement, which is used in [28]. The head position monitor is integrated with the head-mounted display and it is used to measure the yaw (left/right), pitch (up/down), and tilt movements of the head. From these measurements, the eye gaze direction with respect to the virtual scene can be computed accurately. In [28], the social attention pattern of children with ASD is studied entirely using eye gaze as the sole input from the participants with a head-mounted eye tracker.

Because desktop based eye-trackers are fixated to the monitor, eye tracking alone is not sufficient to accurately determine the gaze direction with respect to the computer monitor. Hence, desktop eye trackers typically integrate with head tracking based on face recognition to compensate for head

movement, such as the Eye-Trac 6 desktop eye tracker from Applied Science Laboratories, which is used in [123] in their study.

In [56,60], a similar desktop eye tracking device, called Tobii X120, was used in their study. The device tracks at 120 Hz and allows free head movements within the range of 30 cm by 22 cm by 30 cm (for width, height and depth) and it requires the participant to sit at a distance of 70 cm. The technical specification for Tobii X120 does not state how the head orientation is determined. The authors of [60] designed and implemented an eye tracking application to be used in conjunction with the virtual reality application. The two applications are connected via a high-speed network such that they can determine which and if the participant fixated on one of the regions of interest of the avatar, including forehead, left/right eye, nose, mouth, other face areas, and non-face regions. The same eye tracking device was used in a study with social robots [85]. However, because of the different context where the study wanted to know whether or not the participant was looking at the robot instead of a virtual avatar displayed in a computer monitor, a technician was needed to manually observe the eye gaze and constantly calibrate the system. Automated eye gaze is possible even in this context. However, it would require the participant to wear a head tracker, which was proved to be too uncomfortable for many participants (40% of participants did not want to use it) [124,125].

In [126], another desktop mounted EyeLink 1000 Remote eye tracking system was used in their study. In [127], a remote infrared camera (model D6-HS Remote from Applied Science Laboratories) was used to track the participant's eye gaze.

Zheng et al. reported a system that is capable of automatically detecting head orientation and eye gaze using multiple cameras [128–130]. Based on the detected eye gaze (and engagement level), the system adapts its program content in realtime.

We should note that the trustworthiness of the collected gaze data has rarely been examined carefully with the exception of [131], where the authors studied the raw log files and determined that the quality of recognition is sub-optimal. Future research should pay more attention to the data quality issue, not only for eye gaze data, but all data collected to improve the effectiveness of treatment and to protect the safety of the participants if such data are used in a closed-loop fashion to control the flow of the treatment program.

#### 4.4. Human Motion Tracking via Depth Sensors

Microsoft Kinect and similar depth sensing devices such as Asus Xtion PRO enabled a new way of capturing the gestures as well as facial expressions of the participant [132,133].

In [11], a participant can move the robot to the left or right via the left or right wave gestures. The OpenNI and NITE software packages were used to capture the human motion and to recognize the gestures. In [101], Microsoft Kinect sensor was used to track the participant's movement to determine how closely the participant imitated the pre-scripted and tele-operated movement of the robot. Metrics used include the range of motion of joint angles, response delay, and the distance between two time series representing the movement of the robot and that of the participant.

For studies with social robots, depth cameras are often used to track the presence and the activities of the participant for offline analysis regarding the similarity of the imitation and the original actions performed by the robot, as done in [92,101]. In [92], the pose of the participant and the limbs were first identified using the depth frames, and subsequently the head region in the corresponding color image frames were cropped. From the cropped face, the pose of the head was determined using a computer vision algorithm. From the head pose data extracted from the logged data during the experiment, the joint attention behaviors of children with ASD and typical developing children are compared. The imitation score was determined by the fraction of postured successfully imitated by the participant.

Enabled by inexpensive depth sensing devices such as Microsoft Kinect, it becomes possible to engage children with ASD with motion-based touchless serious games [22,23,25,134]. Unlike other serious games, a participant engages in the game play by moving his entire body. The participant is being continuously tracked by the depth sensor and an avatar representing the participant becomes

part of the game displayed. These types of games not only could be more engaging, but also may enhance the motor skills of the participants too.

Zhang et al. reported their work on using Kinect to track a participant's gesture and assess the quality of the imitation based a set of predefined rules [135]. Feedback is provided in realtime to the participant. Ge et al. [136] used Kinect to track the engagement level of the participant. Yun et al. also used Kinect to track the activity of the participant when he/she is interacting with a social robot [137]. Kinect has also been used to track the movement patterns of the participations [29] and for self-modeling intervention for children with ASD [138].

#### 4.5. Speech Recognition

In some studies, speech recognition is employed to enable a participant to answer questions or engage in a program via predefined voice commands. At this stage, we have yet to see work that has incorporated unscripted conversations.

In [11], a participant can interact with a robot by a number of predefined voice commands, such as move left/right, follow me, stop, and goodbye. The voice commands were captured using the Asus Xtion PRO device, which is equipped with two microphones. The Microsoft Speech Platform Software Development Kit was used to perform speech recognition. In [105], the robot Lucy is said to have speech recognition capability. However, the paper did not provide any technical detail and the use scenarios of the capability in their study.

#### 4.6. Physiological Data Collection

It is well known that physiological characteristics reflect the emotion and attention of a person. Hence, it may be desirable to obtain such information during a session of treatment to gain better insight to the state of an individual with ASD. We have seen more autism research that employs various sensors to collected physiological signals, including Pulse Plethysmogram (PPG), Skin Temperature (ST), Electrodermal Activity (EDA), Electrocardiographic (ECG), and Electroencephalography (EEG) [39,40,51,139,140]. Typically, the classification of such signals was done offline. Hence, the classification accuracy does not impact the flow of the program and cannot cause harm to the participants. Nevertheless, in the long run, obtaining responses through EEG has the potential to develop brain-computer interfaces to improve social skills of individuals with ASD [39]. EEG data were also collected in [141–144]. In [145], Ozcan et al. reported a design for wearable companion toys in which biosensors can be embedded to measure the participant's physiological and emotional state. In [146], physiological data, including PPG, Galvanic Skin Response (GSR), and Respiration (RSP), are collected while a participant with ASD engages in a driving task. Note that in some studies, multimodal data are collected [143]. In addition to physiological data, eye gaze and spoken conversations are also collected [143] to enable adaptive social interaction.

#### 4.7. Video/Voice Recording

Video coding is the standard method for analysis in behavioral research [147,148]. Many autism studies that have been published in clinical journals included video coding, such as [22,38,64,102].

### 5. Program Customizability and Adaptability

Some programs, especially serious games, may allow users to customize their user interface, including the avatar, if it is used, game scene, as well as difficulty level. In [61], a user is asked to customize the avatar and the virtual environment of his/her choice. The user is also asked to calibrate the tactile feedback of the game controller he/she will be using in the program because the authors anticipated that different users would have different sensitivity preferences. Furthermore, different methods were provided in [61] for a user to indicate the specific emotion they recognized, including a visual non-verbal code where each emotion to be recognized was represented by a particular color/icon (happiness was represented by yellow, anger was represented by red, disgust

was represented purple, fear was represented by green, sadness was represented by light blue, surprise was indicated by dark blue, pain was represented by black, neutral was represented by white, and funny face was represented by a trash can), emotional words that told the emotion as it was, and idiomatic expressions, which were short sentences containing a figurative meaning, e.g., an idiomatic expression for fear could be “to get cold feet”. In [25], all three motion-based games were designed for customizability. A user could decide on appropriate game speed, object density, and which body parts to enable for game playing.

Different from program customization, it might be desirable for a program to dynamically adjust its flow of control, task difficulty level, or graphic content based on how a user responded to deliver a personalized program for each user [149]. Of course, this does not apply to all programs because some studies were designed to establish patterns of behaviors of individuals with ASD. In these studies, obviously there is no need to adapt the program flow and content based on the input from the participants. Jarrold et al. [28] studied social attention patterns of high functioning children with ASD in a virtual public speaking task based on eye gaze.

Many treatment programs are not adaptive in that the flow of the programs are not changed dynamically based on the input collected while being administered. For mobile games [14], a participant must provide the correct answer to each question before the participant can complete the current level, and only after completing the current level, can the participant continue on to the next level. Such games do not consider how many times a participant has failed a level.

An adaptive program requires the processing of input data from participants in real time. This can be done either automatically using technology and computer algorithms, or by a therapist manually.

### 5.1. Automatic Adaption of Programs

In [40], the difficulty level is adjusted based on whether or not the participant answered a question correctly via a menu-driven interface. The program has four levels of task difficulty. The difficulty level is reflected from the questions asked for each story. If the answer from the participant is deemed adequate, the next question will be elevated to a higher level unless the current level is already the highest. Mouring and Tang [150] reported an adaptive training for adolescents with ASD using neural network to monitor participant performance and a learning fuzzy inference system to determine deficits to instructional techniques.

Eye tracking is one of the few relatively mature technologies that has been used to retrieve nonverbal communication of a participant in realtime. Some studies used eye gaze information to alter the scene displayed for a participant. In [123], Courgeon et al. studied the joint attention pattern using eye tracking and virtual avatars. The virtual display blurs all the scene except a small rectangular region centered on the point of gaze of the participant. This is done so to provide visual feedback on what the participant is looking at, which enhances self-awareness of gaze.

In [28], the program included a cued condition where an avatar would fade to 70% transparency in six seconds if the participant failed to fixate his/her eyes on the particular avatar when prompted. The entire cued condition is fully automated enabled by the virtual reality program and gaze tracking via a head-mounted display and sensors. The avatar would become opaque again when the participant looked at it again.

An individualized adaptive response module constitutes a core component in the virtual reality system used in [37,41]. The module consists of two adaptive mechanisms, namely, the performance-sensitive system (PS), and the engagement-sensitive system (ES). In PS, the task difficulty is adjusted solely based on the input provided by the participant via the menu-based conversation interface. In ES, the task difficulty is adjusted based on both the input provided by the participant and the engagement level as revealed by the three features obtained from eye gaze tracking data, i.e., the mean pupil diameter, the mean blink rate, and the average fixation duration for each region of interest. For PS, the task difficulty is adjusted based on the performance of the previous task. The performance is determined by the points earned on a task. If the participant earns 70% or more

the maximum score possible, then the performance is regarded as adequate. Otherwise, it is regarded as inadequate. If the performance on a task is deemed adequate, then the difficulty level of the next task is elevated to the next higher level. Otherwise, it is switched to a lower difficulty level. For ES, the rule for adjusting difficulty levels is more complicated because it must fuse both the performance from participant direct input (categorized as adequate and inadequate), and the performance from eye gaze analysis, which is categorized into “good enough” and “not good enough”.

It is encouraging to see that this line of research is gaining momentum. In addition to eye gaze, we have seen studies that employ various methods to obtain user feedback for adaptive control of a treatment/education program, including facial expression, head orientation, human gesture tracking, and physiological data collection [135,136,141,143,146,151].

## 5.2. Manual Control

An adaptive program requires realtime analysis of the responses of a participant during a program in one or more aspects and decides on the most appropriate next step autonomously. To a large degree, there is still a gap towards this goal. Hence, a psychologist or a trained clinician is often required to provide their subjectively evaluation based on their observation during the execution of a program. Sometimes, the professional is asked to not only evaluate the response of the participant, but also to control the flow of the program as well, which is true especially for social robotic based studies, where a social robot is controlled manually by the professional based on his/her judgment. This scheme is often referred to as Wizard of Oz scheme [3,152]. The professional evaluation can also be used to assess the effectiveness of the program.

In Carter et al.’s study of interactive avatars [38], the system output was verbal utterances by a human actor that tele-operated the avatar to talk on the screen. The avatar tried to elicit responses (verbal statements and gestures) from the participants and occasionally called on those who raised their hands (and their parents) to engage in one-on-one conversation. In [64], a carton-like avatar was controlled by an actual person. The motivation for doing so is to gain more flexibility during the interactive session because the current state-of-the-art technology is still incapable of handling spontaneous conversations.

In [85], a technician was used to monitor the eye gaze of the participant both for the purpose of calibration of the eye tracking system and for identifying when the participant looked at the robot and the object as directed by the robot. When the participant performed as instructed within 7 s, the technician also triggered a reward for the participant by pushing a button in the system. Otherwise, a miss event was recorded and the next task was administered.

In Dehkordi et al.’s study [106], the system output consists of simple motions of the robot parrot following the commands of the operator, as well as voice from the speaker uttered by the operator. This involves the use of body motor to control the motions of the wings, the legs, and the neck, and the use of head motor to control the motions of the eyes and the beak. The operator (ideally a therapist) can use the camera on the cage to check the validity of the execution of commands sent to the robot parrot, use the speaker and microphone to verbally communicate with the ASD child, and adjust both the digital commands and verbal output based on what the child says or does.

In [95], the robots were also controlled by a trainer using a laptop running a custom program. The trainer is responsible for deciding when to greet a participant, warming up a game with body stretching activities, playing various games, and deciding when to say goodbye. The same scheme was used in [96].

In motion-based touchless serious games [22,25], a therapist was also tasked to observe the behaviors of participant and to intervene when necessary.

## 6. Program Evaluation

In addition to program adaptivity, objective evaluation of a program would also need to analyze the input data collected during a study. There are different types of program evaluations.

For a treatment program, the evaluation focuses on its impact on the participants, which typically relies on traditional clinical-proven standard tests, but maybe assessed via technology-based methods as well. For a program aimed to diagnose autism, or a program aimed to establish patterns on social behaviors of individuals with ASD, the purpose of program evaluation is to establish the reliability of the patterns detected or diagnosis [60]. Yet some studies focus on the evaluation of effectiveness of different ways of delivering a treatment plan to individuals with ASD such as [38]. Many pilot studies simply validated a predefined set of hypotheses using their programs [28,38,41,85]. In all studies, the usability of the programs used could also be evaluated [61]. A special case for usability study is testing the acceptability of a new way of delivering content to individuals with ASD, such as [48,153], where a consumer-grade head-mounted virtual reality goggle was experimented.

### 6.1. Behavioral Pattern Assessment

In [60], the authors studied the patterns of facial emotion recognition and related characteristics for both children with ASD and typical developing children using a virtual avatar based program and eye tracking. They used a number of metrics, including the recognition accuracy of facial emotions, the response time and the confidence level of making the recognition. The authors found that it takes longer for children with ASD to make a recognition with lower confidence level than that of typical developing children. Furthermore, children with ASD have more variation in eye gaze patterns than those of typical developing children. However, the emotion recognition rate is rather similar between the two groups.

### 6.2. Evaluation of Treatment Effectiveness

In [53], emotion recognition skills of the participants were assessed using six Ekman and Friesen photographs [154] and six drawings of facial expressions, facial recognition skills were evaluated using the Benton Facial Recognition Test [155], social interaction skills in natural environments were rated by the Social Skills Rating System [156] and social skills observation. To validate that a treatment program is effective, these tests are typically administered before and after the treatment sessions and the results are compared.

In a sequence of studies [95–97], Srinivasan et al. presented a systematic evaluation of three different ten-week intervention programs with pretest and posttest using standardized clinical tests. In the first study [95], the repetitive maladaptive behaviors as well as the affective states of participants were coded by clinicians and then compared using statistical analysis. In the second study [96], a standardized test of joint attention was used to determine the verbal communication skills of the participants before and after the ten-week intervention program. In the third study [97], a standardized test of motor performance called BOT-2 was used in the pretest and posttest to assess the effectiveness of the intervention program on imitation, interpersonal synchrony, and motor performance of children with ASD. In all these studies, the authors observed that there were no advantages of using social robots compared with rhythm based training.

In [14], a clinician was asked to evaluate the following aspects of the game: (1) how well the game was able to attract the attention of the participant; (2) how well the participant was able to carry out tasks on their own; (3) the degree of effort the participant had made to find the right answer; (4) how well the participant showed signs of anticipation; (5) how well the participant was able to wait patiently to transition from one level to another; and (6) the degree of self-esteem had shown by the participant during the game with respect to the confidence displayed when making successful attempts and despite failures.

In [61], the effectiveness of the treatment program using JeStiMule was evaluated by comparing the results of the facial emotion recognition accuracy obtained before the four-week long treatment program and after using both the avatars employed in JeStiMule and the pictures of real-life characters. They observed statistically significant improvement of the treatment plan.

To assess the effectiveness of using motion-based touchless serious games to improve autistic children's learning, two variables, namely selective attention (i.e., how well a participant could focus on an important stimulus in the presence of distractions) and sustained attention (i.e., how well a participant could hold his/her attention to complete a task), were measured and compared using the Bell Test at the beginning, during, and seven days after the end of the treatment [22]. Data for the two variables were extracted from the recorded videos taken during the game plays. The researchers noticed an increase of the two variables for all children. In a follow-up study [25], assessment was made at the beginning and the end of treatment period. In addition to select attention and sustained attention, three additional variables were used, including visual perception (i.e., how well a participant could process visual information), motor coordination (i.e., how well a participant could move body parts coherently), and visuo-motor integration (i.e., how well a participant could control his/her body movement when prompted visually).

In some studies, both parameters collected while a participant was playing a serious game (or a sequence of games) and standardized tests were used to assess the effectiveness of a treatment plan. The data were collected at the beginning and at the end of the six-month treatment plan. In [13], the game-associated parameters used for assessment include the game identifier and level of difficulty, when the game was played, the game score, and the reaction time for a participant to complete a task. The clinical measures include the Autism Diagnosis Observation Schedule, the Vineland Adaptive Behavior Scale II, Weschsler scales, the Child Behavior Checklist, and the Social Communication Questionnaire, and the Parenting Stress Index. The researchers found that computer-based serious games benefited children with ASD. In particular, they found that 4 out of 6 imitation games significantly improved the participants' imitation scores, and 3 out of 4 joint attention games shortened the time it took for participants to complete the task over multiple sessions.

In [19], the Social Responsiveness Scale, 2nd Edition (SRS-2) standardized test was used in pretest and posttest to determine the effectiveness of a serious game called emotiplay.

### 6.3. Usability Assessment

In [61], the usability of the program called JeStiMule was evaluated in terms of its adaptability, effectiveness and efficiency with a group of individuals with ASD. For adaptability, the authors observed that all participants managed to use the gamepad to control the avatar, 91% of participants understood the association between facial emotion and the corresponding non-verbal code, and different participants chose their own favorite ways of indicating the facial emotion that they have identified. The effectiveness is determined by the percentage of participants who managed to complete different modules in JeStiMule. 73% of participants were able to complete all modules. The efficiency is determined by how long the participants were able to complete each module among those who have completed all modules. On average, it takes 49 minutes for a participant to complete each module.

In [48], the acceptability and immersiveness of the consumer-grade head-mounted goggle, Oculus Rift, among individuals with ASD were examined via a two-phase study. The acceptability of the device is determined by the fraction of participants who managed to complete the study. Out of 29 participants, 25 completed the phase I study. 23 participants proceeded on to complete the second phase of the study. For the immersiveness of the device, the authors chose to use the Independent Television Commission-Sense of Presence Inventory (ITC-SoPI) to evaluate the subjective experiences by participants who have completed the study. ITC-SoPI evaluates user experiences in the following four categories: (1) spatial presence, which is about how a participant feels he/she is really present in the physical environment; (2) user engagement, which is about how involved a participant is in a virtual reality program and how intense it is; (3) ecological validity, which is about how natural a participant feels about the experience; and (4) negative effects, which are about whether or not a participant feels any dizziness, headache, etc. This study also briefly analyzed participant behavior while using the device by first video taping the sessions and then analyzing the recorded video qualitatively manually.

## 7. Discussion

As we have seen in the reviewed literature, there are two primary means of eliciting proper reactions from individuals with ASD, one using virtual reality technologies, and the other using social robotic technologies. The former typically relies on one or more avatars to serve as an agent to communicate with a participant, while the latter uses a robot. Virtual reality has two advantages over robotics: (1) both the avatar and the virtual scenes can be changed dynamically that fit the treatment scenarios, while a robot cannot easily change its appearance and the treatment environments cannot be changed easily; (2) the virtual reality based systems can be much cheaper than robots. Hence, unless the cost of robots is reduced drastically, virtually reality based systems could be more attractive to home-based uses.

That said, robotic based solutions provide physical embodiment, which is not possible for virtual reality based systems. This advantage may lead to better generalization of skills learned in treatment sessions to the real-world scenarios, although this speculation has yet to be proven in clinical trials.

For virtual reality based systems, even though it is desirable to render 3D views to a participant, which could potentially increase the treatment effectiveness and transferability to real-world skills, the use of head-mounted displays or virtual reality goggles might not be suitable for all individuals with ASD because some are more prone to cyber sickness. This concern is also applicable for wearable devices that take physiological data from a participant.

Another observation is that the current autism studies have predominately focused on treatment programs that target a specific set of social skills. Although temporary improvements in social skills for individuals with ASD have been observed, it is doubtful such improvements can last and can be generalized. More importantly, such training programs do not address the psychological and emotion issues for individuals (especial for children and adolescents) with ASD. What is urgently needed for an individual with ASD is a life coach who he/she trusts, who shares the same interest with him/her with, who knows his/her emotions, and who is never upset no matter how badly he/she behaves. A recently granted patent describes a system that partially fulfill this goal [157]. This system provides an avatar-based mobile interface for an autism child to bound with the avatar character and to learn various lessons on social and emotional regulation skills. The system relies on the use of one or more therapist or parent to interact with an autism child. Essentially, the system presents an illusion that the child with ASD is interacting with an intelligent avatar who he/she feels comfortable interacting with while in reality he/she is communicating with an actual human being who he/she otherwise would be reluctant to communicate with. To help the participating parent or therapist, the system also provides carefully crafted prompts to guide him/her on how to properly interact with the child with ASD. We are currently also working on a system that aims to serve as a virtual life coach for children with ASD [158].

Several publications have reported so called closed-loop autonomous systems based on social robots or virtual avatars [135]. Unfortunately, such systems have achieved autonomy only for very specific scenarios. For example, in [135], the system is capable of assessing a very small set of imitation gestures made by a participant in realtime using predefined rules. This is far from the degree of autonomy needed to train children with ASD towards improved practical social skills. In [159], an ambitious design towards a virtual avatar that can autonomously respond to the participant's emotional reactions and action tendencies was presented. Unfortunately, the authors conceded that it is exceedingly difficult to model the participant accurately based on the limited information collected, including the frequency of touch (on the screen) and eye tracking. In the end, they supplemented their system with a Wizard of Oz control panel for a therapist to intervene when needed.

A much more viable approach is proposed by Esteban et al. where the design goal should be a supervised autonomous system [160]. According to this design, autonomy can be achieved for specific tasks, such as to determine where a social robot (or a virtual avatar) looks at next. With supervised autonomy, a therapist is needed to provide high level guidance and if necessary, to override the decision made autonomously. To accomplish this goal, the system would need to track several primary and

secondary variables. The primary variables include the performance of the participant who performs the designated tasks. The secondary variables include: (1) the level of social engagement reflected by eye-contact and verbal utterances; (2) the emotional level (positive or negative); (3) the behavioral level (stereotypical, adaptive, and maladaptive); (4) the cognitive level (rational or irrational). In addition to tracking the participant, the system proposed by Esteban et al. also tracks the behavior of the social robot for maximum safety of the participant [160]. In this system, three regular cameras and three Kinect sensors are used (one of the Kinect sensor is used to self-monitor the social robot).

With the supervised autonomous system design, we believe machine learning and artificial intelligence [161–165], in conjunction with multimodal sensing, will play much more important roles in the next generation autism diagnosis and treatment programs. Considering that good therapists are always in very high demand for the autism population, we envisage that a therapist could operate multiple supervised autonomous systems remotely at different sites, which could benefit more children with ASD while reducing the cost of autism care.

## 8. Conclusions

In this article, we reviewed the autism research facilitated by various technologies from an engineering perspective. We primarily focused on research results produced in the last five years (2012–2017). As can be seen, while significant progress has been made, there are major challenges to overcome in future research. First, other than manual input via keyboard/mouse/touch-screen, the only mature technology for automatically capturing participant responses is eye gaze. To assess how well a participant responded to an intervention program, current research predominately relied on video tapping and offline analysis of the recorded videos. Even though there have been efforts to use various motion/facial tracking and physiological tracking devices to capture the emotional state of the participant, such efforts are still in preliminary stages. Second, few autism treatment programs are customizable and automatically adaptive. This is obviously a limitation for autism studies because individuals with ASD are known to be highly different in their deficiencies, skill levels and preferences. To overcome this limitation, some studies relied on the use of therapists to operate the virtual avatar or social robot in realtime. However, this approach is not only not conducive for reducing the cost of autism treatment using technology, but also has negative impact on the repeatability of experiments. Furthermore, the quality of the program control also heavily depends on the training and expertise of the therapist. Hence, the solution is not scalable. The resolving of the first issue will open the door for designing and implementing fully autonomous virtual avatars and social robots that help individuals with ASD.

Finally, we provide our observation on how technological limitations might be affecting ASD therapies. First, there lacks a highly reliable and accurate sensing platform that can capture and recognize the affective state of the participant. As pointed in [160], this is the prerequisite to achieving autonomous control of a therapy program that is personalized for each individual and is capable of dynamically adapting the program content based on the affective state of the individual. Many current studies failed to consider the affective state of the participants completely. Even for those that do, the technologies employed are ad-hoc, are often without establishing the ground-truth, and are not processed in realtime.

Second, the therapy program content is typically designed one-size-for-all. It is well known that children with ASD typically have intense interests in certain topics. The therapy programs could be much more effective to an individual if the program content aligns well with the individual's intense interests. Getting to know which participant to have what intense interests would go beyond technology. Ideally, for each child with ASD, an individualized educational placement (IEP) team would be needed to provide intimate knowledge about the child so that the content of the therapy program can be customized for the individual [158].

Third, there lack longitudinal studies on the long-term effectiveness of technology-based therapies. While it is accepted that children with ASD are generally more receptive to technology, for example,

they show less anxiety when interacting with virtual avatars or social robots, it is unclear such technologies can truly help children with ASD to gain social skills and emotional regulation skills in practice in the long-run when they have to directly interact with real human beings day-in and day-out.

**Acknowledgments:** This article was supported in part by the Cleveland State University Graduate Faculty Research Support Program, and by the Cleveland State University Faculty Research and Development—Internet of Things Program. The authors did not receive any fund for covering the costs to publish in open access.

**Author Contributions:** Xiongyi Liu and Wenbing Zhao conceived and designed the review protocol, Xiongyi Liu, Qing Wu, Wenbing Zhao, and Xiong Luo performed literature search and reviewed the papers collected, and Xiongyi Liu, Qing Wu, Wenbing Zhao and Xiong Luo wrote the paper.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

## Abbreviations

The following abbreviations are used in this manuscript:

ASD	Autism Spectrum Disorder
DOF	Degrees of Freedom
EKG	Electrocardiographic
EDA	Electrodermal Activity
EEG	Electroencephalography
ES	Engagement-Sensitive System
GSR	Galvanic skin response
ITC-SoPI	Independent Television Commission-Sense of Presence Inventory
PPG	Pulse Plethysmogram
RSP	Respiration
SKT	Skin Temperature
PS	Performance-Sensitive System

## References

1. Wieckowski, A.T.; White, S.W. Application of technology to social communication impairment in childhood and adolescence. *Neurosci. Biobehav. Rev.* **2017**, doi:10.1016/j.neubiorev.2016.12.030.
2. Aresti-Bartolome, N.; Garcia-Zapirain, B. Technologies as support tools for persons with autistic spectrum disorder: A systematic review. *Int. J. Environ. Res. Public Health* **2014**, *11*, 7767–7802.
3. Scassellati, B.; Admoni, H.; Matarić, M. Robots for use in autism research. *Ann. Rev. Biomed. Eng.* **2012**, *14*, 275–294.
4. Diehl, J.J.; Schmitt, L.M.; Villano, M.; Crowell, C.R. The clinical use of robots for individuals with autism spectrum disorders: A critical review. *Res. Autism Spectr. Disord.* **2012**, *6*, 249–262.
5. Cabibihan, J.J.; Javed, H.; Ang, M.; Aljunied, S.M. Why robots? A survey on the roles and benefits of social robots in the therapy of children with autism. *Int. J. Soc. Robot.* **2013**, *5*, 593–618.
6. Pennisi, P.; Tonacci, A.; Tartarisco, G.; Billeci, L.; Ruta, L.; Gangemi, S.; Pioggia, G. Autism and social robotics: A systematic review. *Autism Res.* **2015**, *9*, 165–183.
7. Begum, M.; Serna, R.W.; Yanco, H.A. Are robots ready to deliver autism interventions? a comprehensive review. *Int. J. Soc. Robot.* **2016**, *8*, 157–181.
8. Parsons, S. Authenticity in Virtual Reality for assessment and intervention in autism: A conceptual review. *Educ. Res. Rev.* **2016**, *19*, 138–157.
9. Zakari, H.M.; Ma, M.; Simmons, D. A review of serious games for children with autism spectrum disorders (ASD). In Proceedings of the International Conference on Serious Games Development and Applications, Berlin, Germany, 9–10 October 2014; Springer: Cham, Switzerland, 2014; pp. 93–106.
10. Grossard, C.; Grynspan, O.; Serret, S.; Jouen, A.L.; Bailly, K.; Cohen, D. Serious games to teach social interactions and emotions to individuals with autism spectrum disorders (ASD). *Comput. Educ.* **2017**, *113*, 195–211.

11. Chuah, M.C.; Coombe, D.; Garman, C.; Guerrero, C.; Spletzer, J. Lehigh instrument for learning interaction (lili): An interactive robot to aid development of social skills for autistic children. In Proceedings of the IEEE 11th International Conference on Mobile Ad Hoc and Sensor Systems, Philadelphia, PA, USA, 28–30 October 2014; pp. 731–736.
12. Marwecki, S.; Rädle, R.; Reiterer, H. Encouraging Collaboration in Hybrid Therapy Games for Autistic Children. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*; ACM: New York, NY, USA, 2013; pp. 469–474.
13. Jouen, A.L.; Narzisi, A.; Xavier, J.; Tilmont, E.; Bodeau, N.; Bono, V.; Ketem-Premel, N.; Anzalone, S.; Maharatna, K.; Chetouani, M.; et al. GOLIAH (Gaming Open Library for Intervention in Autism at Home): A 6-month single blind matched controlled exploratory study. *Child Adolesc. Psychiatry Ment. Health* **2017**, *11*, 17.
14. Chua, L.; Goh, J.; Nay, Z.T.; Huang, L.; Cai, Y.; Seah, R. ICT-Enabled Emotional Learning for Special Needs Education. In *Simulation and Serious Games for Education*; Springer: Cham, Switzerland, 2017; pp. 29–45.
15. Zhang, L.; Gabriel-King, M.; Armento, Z.; Baer, M.; Fu, Q.; Zhao, H.; Swanson, A.; Sarkar, M.; Warren, Z.; Sarkar, N. Design of a Mobile Collaborative Virtual Environment for Autism Intervention. In Proceedings of the International Conference on Universal Access in Human-Computer Interaction, Toronto, ON, Canada, 17–21 July 2016; Springer: Cham, Switzerland, 2016; pp. 265–275.
16. Vullamparthi, A.J.; Nelaturu, S.C.B.; Mallaya, D.D.; Chandrasekhar, S. Assistive learning for children with autism using augmented reality. In Proceedings of the 2013 IEEE Fifth International Conference on Technology for Education (T4E), Kharagpur, India, 18–20 December 2013; pp. 43–46.
17. Simões, M.; Mouga, S.; Pedrosa, F.; Carvalho, P.; Oliveira, G.; Branco, M.C. Neurohab: a platform for virtual training of daily living skills in autism spectrum disorder. *Procedia Technol.* **2014**, *16*, 1417–1423.
18. Aresti-Bartolome, N.; Garcia-Zapirain, B. Cognitive rehabilitation system for children with autism spectrum disorder using serious games: A pilot study. *Bio-Med. Mater. Eng.* **2015**, *26*, S811–S824.
19. Fridenson-Hayo, S.; Berggren, S.; Lassalle, A.; Tal, S.; Pigat, D.; Meir-Goren, N.; O'Reilly, H.; Ben-Zur, S.; Bölte, S.; Baron-Cohen, S. Emotiplay: A serious game for learning about emotions in children with autism: Results of a cross-cultural evaluation. *Eur. Child Adolesc. Psychiatry* **2017**, *26*, 979–992.
20. Zhao, H.; Swanson, A.; Weitlauf, A.; Warren, Z.; Sarkar, N. A Novel Collaborative Virtual Reality Game for Children with ASD to Foster Social Interaction. In Proceedings of the International Conference on Universal Access in Human-Computer Interaction, Toronto, ON, Canada, 17–21 July 2016; Springer: Cham, Switzerland, 2016; pp. 276–288.
21. Bono, V.; Narzisi, A.; Jouen, A.L.; Tilmont, E.; Hommel, S.; Jamal, W.; Xavier, J.; Billeci, L.; Maharatna, K.; Wald, M.; et al. GOLIAH: A gaming platform for home-based intervention in autism—principles and design. *Front. Psychiatry* **2016**, *7*, 70.
22. Bartoli, L.; Corradi, C.; Garzotto, F.; Valoriani, M. Exploring motion-based touchless games for autistic children's learning. In Proceedings of the 12th International Conference on Interaction Design and Children, New York, NY, USA, 24–27 June 2013; pp. 102–111.
23. Garzotto, F.; Gelsomini, M.; Oliveto, L.; Valoriani, M. Motion-based touchless interaction for ASD children: A case study. In Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces, Como, Italy, 27–29 May 2014; pp. 117–120.
24. Ge, Z.; Fan, L. Social Development for Children with Autism Using Kinect Gesture Games: A Case Study in Suzhou Industrial Park Renai School. In *Simulation and Serious Games for Education*; Springer: Cham, Switzerland, 2017; pp. 113–123.
25. Bartoli, L.; Garzotto, F.; Gelsomini, M.; Oliveto, L.; Valoriani, M. Designing and Evaluating Touchless Playful Interaction for ASD Children. In Proceedings of the 2014 Conference on Interaction Design and Children, Aarhus, Denmark, 17–20 June 2014; ACM: New York, NY, USA, 2014; pp. 17–26.
26. Barajas, A.O.; Al Osman, H.; Shirmohammadi, S. A Serious Game for children with Autism Spectrum Disorder as a tool for play therapy. In Proceedings of the 2017 IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH), Perth, WA, Australia, 2–4 April 2017; pp. 1–7.
27. Whyte, E.M.; Smyth, J.M.; Scherf, K.S. Designing serious game interventions for individuals with autism. *J. Autism Dev. Disord.* **2015**, *45*, 3820–3831.

28. Jarrold, W.; Mundy, P.; Gwaltney, M.; Bailenson, J.; Hatt, N.; McIntyre, N.; Kim, K.; Solomon, M.; Novotny, S.; Swain, L. Social attention in a virtual public speaking task in higher functioning children with autism. *Autism Res.* **2013**, *6*, 393–410.
29. Ip, H.H.; Lai, C.H.Y.; Wong, S.W.; Tsui, J.K.; Li, R.C.; Lau, K.S.Y.; Chan, D.F. Visuospatial attention in children with Autism Spectrum Disorder: A comparison between 2-D and 3-D environments. *Cogent Educ.* **2017**, *4*, 1307709.
30. Lorenzo, G.; Lledó, A.; Pomares, J.; Roig, R. Design and application of an immersive virtual reality system to enhance emotional skills for children with autism spectrum disorders. *Comput. Educ.* **2016**, *98*, 192–205.
31. Bozgeyikli, L.; Bozgeyikli, E.; Raij, A.; Alqasemi, R.; Katkooori, S.; Dubey, R. Vocational training with immersive virtual reality for individuals with autism: Towards better design practices. In Proceedings of the 2016 IEEE 2nd Workshop on Everyday Virtual Reality (WEVR), Greenville, SC, USA, 20 March 2016; pp. 21–25.
32. Bozgeyikli, E.; Raij, A.; Katkooori, S.; Dubey, R. Locomotion in Virtual Reality for Individuals with Autism Spectrum Disorder. In Proceedings of the 2016 Symposium on Spatial User Interaction, Tokyo, Japan, 15–16 October 2016; pp. 33–42.
33. Ip, H.H.; Wong, S.W.; Chan, D.F.; Byrne, J.; Li, C.; Yuan, V.S.; Lau, K.S.; Wong, J.Y. Virtual reality enabled training for social adaptation in inclusive education settings for school-aged children with autism spectrum disorder (ASD). In Proceedings of the International Conference on Blending Learning, Beijing, China, 19–21 July 2016; Springer: Cham, Switzerland, 2016; pp. 94–102.
34. Cheng, Y.; Huang, C.L.; Yang, C.S. Using a 3D immersive virtual environment system to enhance social understanding and social skills for children with autism spectrum disorders. *Focus Autism Other Dev. Disabil.* **2015**, *30*, 222–236.
35. Maskey, M.; Lowry, J.; Rodgers, J.; McConachie, H.; Parr, J.R. Reducing specific phobia/fear in young people with autism spectrum disorders (ASDs) through a virtual reality environment intervention. *PLoS ONE* **2014**, *9*, e100374.
36. Bozgeyikli, L.; Bozgeyikli, E.; Clevenger, M.; Gong, S.; Raij, A.; Alqasemi, R.; Sundarrao, S.; Dubey, R. VR4VR: Towards vocational rehabilitation of individuals with disabilities in immersive virtual reality environments. In Proceedings of the 2014 2nd Workshop on Virtual and Augmented Assistive Technology (VAAT), Minneapolis, MN, USA, 30 March 2014; pp. 29–34.
37. Lahiri, U.; Bekele, E.; Dohrmann, E.; Warren, Z.; Sarkar, N. Design of a virtual reality based adaptive response technology for children with autism. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2013**, *21*, 55–64.
38. Carter, E.J.; Williams, D.L.; Hodgins, J.K.; Lehman, J.F. Are children with autism more responsive to animated characters? A study of interactions with humans and human-controlled avatars. *J. Autism Dev. Disord.* **2014**, *44*, 2475–2485.
39. Amaral, C.P.; Simões, M.A.; Castelo-Branco, M.S. Neural Signals Evoked by Stimuli of Increasing Social Scene Complexity Are Detectable at the Single-Trial Level and Right Lateralized. *PLoS ONE* **2015**, *10*, e0121970.
40. Kuriakose, S.; Lahiri, U. Understanding the psycho-physiological implications of interaction with a virtual reality-based system in adolescents with autism: a feasibility study. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2015**, *23*, 665–675.
41. Lahiri, U.; Bekele, E.; Dohrmann, E.; Warren, Z.; Sarkar, N. A physiologically informed virtual reality based social communication system for individuals with autism. *J. Autism Dev. Disord.* **2015**, *45*, 919–931.
42. Yang, Y.D.; Allen, T.; Abdullahi, S.M.; Pelphrey, K.A.; Volkmar, F.R.; Chapman, S.B. Brain responses to biological motion predict treatment outcome in young adults with autism receiving Virtual Reality Social Cognition Training: Preliminary findings. *Behav. Res. Ther.* **2017**, *93*, 55–66.
43. Wallace, S.; Parsons, S.; Bailey, A. Self-reported sense of presence and responses to social stimuli by adolescents with ASD in a collaborative virtual reality environment. *J. Intellect. Dev. Disabil.* **2017**, *42*, 131–141.
44. Forbes, P.A.; Pan, X.; Hamilton, A.F. Reduced mimicry to virtual reality avatars in Autism Spectrum Disorder. *J. Autism Dev. Disord.* **2016**, *46*, 3788–3797.
45. Didehbani, N.; Allen, T.; Kandalaf, M.; Krawczyk, D.; Chapman, S. Virtual reality social cognition training for children with high functioning autism. *Comput. Hum. Behav.* **2016**, *62*, 703–711.
46. Ke, F.; Im, T. Virtual-reality-based social interaction training for children with high-functioning autism. *J. Educ. Res.* **2013**, *106*, 441–461.

47. Wang, M.; Reid, D. Using the virtual reality-cognitive rehabilitation approach to improve contextual processing in children with autism. *Sci. World J.* **2013**, *2013*, 716890.
48. Newbutt, N.; Sung, C.; Kuo, H.J.; Leahy, M.J.; Lin, C.C.; Tong, B. Brief report: A pilot study of the use of a virtual reality headset in autism populations. *J. Autism Dev. Disord.* **2016**, *46*, 3166–3176.
49. Miller, H.L.; Bugnariu, N.L. Level of immersion in virtual environments impacts the ability to assess and teach social skills in autism spectrum disorder. *Cyberpsychol. Behav. Soc. Netw.* **2016**, *19*, 246–256.
50. Wang, X.; Desalvo, N.; Gao, Z.; Zhao, X.; Lerman, D.C.; Gnawali, O.; Shi, W. Eye Contact Conditioning in Autistic Children Using Virtual Reality Technology. In Proceedings of the International Symposium on Pervasive Computing Paradigms for Mental Health, Tokyo, Japan, 8–9 May 2014; Springer: Cham, Switzerland, 2014; pp. 79–89.
51. Li, Y.; Elmaghraby, A.S.; El-Baz, A.; Sokhadze, E.M. Using physiological signal analysis to design affective VR games. In Proceedings of the 2015 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Abu Dhabi, UAE, 7–10 December 2015; pp. 57–62.
52. Carter, E.J.; Hyde, J.; Williams, D.L.; Hodgins, J.K. 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, San Jose, CA, USA, 7–12 May 2016; pp. 140–151.
53. Hopkins, I.M.; Gower, M.W.; Perez, T.A.; Smith, D.S.; Amthor, F.R.; Casey Wimsatt, F.; Biasini, F.J. Avatar assistant: improving social skills in students with an ASD through a computer-based intervention. *J. Autism Dev. Disord.* **2011**, *41*, 1543–1555.
54. Cai, Y.; Chiew, R.; Fan, L.; Kwek, M.K.; Goei, S.L. The Virtual Pink Dolphins Project: An International Effort for Children with ASD in Special Needs Education. In *Simulation and Serious Games for Education*; Springer: Cham, Switzerland, 2017; pp. 1–11.
55. Cai, Y.; Chia, N.K.; Thalmann, D.; Kee, N.K.; Zheng, J.; Thalmann, N.M. Design and development of a virtual dolphinarium for children with autism. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2013**, *21*, 208–217.
56. Feng, Y.; Cai, Y. A Gaze Tracking System for Children with Autism Spectrum Disorders. In *Simulation and Serious Games for Education*; Springer: Cham, Switzerland, 2017; pp. 137–145.
57. Kuriakose, S.; Lahiri, U. Design of a Physiology-Sensitive VR-Based Social Communication Platform for Children With Autism. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, *25*, 1180–1191.
58. Halabi, O.; El-Seoud, S.A.; Alja'am, J.M.; Alpona, H.; Al-Hemadi, M.; Al-Hassan, D. Design of Immersive Virtual Reality System to Improve Communication Skills in Individuals with Autism. *Int. J. Emerg. Technol. Learn.* **2017**, *12*, doi:10.3991/ijim.v11i2.6555.
59. Kim, K.; Rosenthal, M.Z.; Gwaltney, M.; Jarrold, W.; Hatt, N.; McIntyre, N.; Swain, L.; Solomon, M.; Mundy, P. A virtual joy-stick study of emotional responses and social motivation in children with autism spectrum disorder. *J. Autism Dev. Disord.* **2015**, *45*, 3891–3899.
60. Bekele, E.; Crittendon, J.; Zheng, Z.; Swanson, A.; Weitlauf, A.; Warren, Z.; Sarkar, N. Assessing the utility of a virtual environment for enhancing facial affect recognition in adolescents with autism. *J. Autism Dev. Disord.* **2014**, *44*, 1641–1650.
61. Serret, S.; Hun, S.; Iakimova, G.; Lozada, J.; Anastassova, M.; Santos, A.; Vesperini, S.; Askenazy, F. Facing the challenge of teaching emotions to individuals with low-and high-functioning autism using a new Serious game: a pilot study. *Mol. Autism* **2014**, *5*, 37.
62. Blain, S.D.; Peterman, J.S.; Park, S. Subtle cues missed: Impaired perception of emotion from gait in relation to schizotypy and autism spectrum traits. *Schizophr. Res.* **2017**, *183*, 157–160.
63. Roether, C.L.; Omlor, L.; Christensen, A.; Giese, M.A. Critical features for the perception of emotion from gait. *J. Vis.* **2009**, *9*, 15.
64. Hsu, C.W.; Teoh, Y.S. Investigating Event Memory in Children with Autism Spectrum Disorder: Effects of a Computer-Mediated Interview. *J. Autism Dev. Disord.* **2017**, *47*, 359–372.
65. Kandalaft, M.R.; Didehbani, N.; Krawczyk, D.C.; Allen, T.T.; Chapman, S.B. Virtual reality social cognition training for young adults with high-functioning autism. *J. Autism Dev. Disord.* **2013**, *43*, 34–44.
66. Li, C.; Jia, Q.; Feng, Y. Human-Robot Interactoin Design for Robot-Assisted Intervention for Children with Autism Based on ES Theory. In Proceedings of the 2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), Hangzhou, China, 27–28 August 2016; Volume 2, pp. 320–324.

67. Good, J.; Good, J.; Parsons, S.; Parsons, S.; Yuill, N.; Yuill, N.; Brosnan, M.; Brosnan, M. Virtual reality and robots for autism: moving beyond the screen. *J. Assist. Technol.* **2016**, *10*, 211–216.
68. De Haas, M.; Aroyo, A.M.; Barakova, E.; Haselager, W.; Smeekens, I. The effect of a semi-autonomous robot on children. In Proceedings of the 2016 IEEE 8th International Conference on Intelligent Systems (IS), Sofia, Bulgaria, 4–6 September 2016; pp. 376–381.
69. Bernardo, B.; Alves-Oliveira, P.; Santos, M.G.; Melo, F.S.; Paiva, A. An Interactive Tangram Game for Children with Autism. In Proceedings of the International Conference on Intelligent Virtual Agents, Los Angeles, CA, USA, 20–23 September 2016; Springer: Cham, Switzerland, 2016; pp. 500–504.
70. Hong, T.S.; Mohamaddan, S.; Shazali, S.T.S.; Mohtadzar, N.A.A.; Bakar, R.A. A review on assistive tools for autistic patients. In Proceedings of the 2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, Malaysia, 4–8 December 2016; pp. 51–56.
71. Mengoni, S.E.; Irvine, K.; Thakur, D.; Barton, G.; Dautenhahn, K.; Guldborg, K.; Robins, B.; Wellsted, D.; Sharma, S. Feasibility study of a randomised controlled trial to investigate the effectiveness of using a humanoid robot to improve the social skills of children with autism spectrum disorder (Kaspar RCT): A study protocol. *BMJ Open* **2017**, *7*, e017376.
72. Sartorato, F.; Przybylowski, L.; Sarko, D.K. Improving therapeutic outcomes in autism spectrum disorders: Enhancing social communication and sensory processing through the use of interactive robots. *J. Psychiatr. Res.* **2017**, *90*, 1–11.
73. Wong, H.; Zhong, Z. Assessment of robot training for social cognitive learning. In Proceedings of the 2016 16th International Conference on Control, Automation and Systems (ICCAS), Gyeongju, Korea, 16–19 October 2016; pp. 893–898.
74. So, W.C.; Wong, M.Y.; Cabibihan, J.J.; Lam, C.Y.; Chan, R.Y.; Qian, H.H. Using robot animation to promote gestural skills in children with autism spectrum disorders. *J. Comput. Assist. Learn.* **2016**, *32*, 632–646.
75. Bharatharaj, J.; Huang, L.; Al-Jumaily, A.M.; Krageloh, C.; Elara, M.R. Effects of Adapted Model-Rival Method and parrot-inspired robot in improving learning and social interaction among children with autism. In Proceedings of the 2016 International Conference on Robotics and Automation for Humanitarian Applications (RAHA), Kollam, India, 18–20 December 2016; pp. 1–5.
76. Tennyson, M.F.; Kuester, D.A.; Casteel, J.; Nikolopoulos, C. Accessible robots for improving social skills of individuals with autism. *J. Artif. Intell. Soft Comput. Res.* **2016**, *6*, 267–277.
77. Suzuki, R.; Lee, J. Robot-play therapy for improving prosocial behaviours in children with Autism Spectrum Disorders. In Proceedings of the 2016 International Symposium on Micro-NanoMechatronics and Human Science (MHS), Nagoya, Japan, 28–30 November 2016; pp. 1–5.
78. Wainer, J.; Dautenhahn, K.; Robins, B.; Amirabdollahian, F. A pilot study with a novel setup for collaborative play of the humanoid robot KASPAR with children with autism. *Int. J. Soc. Robot.* **2014**, *6*, 45–65.
79. Wainer, J.; Robins, B.; Amirabdollahian, F.; Dautenhahn, K. Using the humanoid robot KASPAR to autonomously play triadic games and facilitate collaborative play among children with autism. *IEEE Trans. Auton. Ment. Dev.* **2014**, *6*, 183–199.
80. Vanderborght, B.; Simut, R.; Saldien, J.; Pop, C.; Rusu, A.S.; Pintea, S.; Lefeber, D.; David, D.O. Using the social robot probo as a social story telling agent for children with ASD. *Interact. Stud.* **2012**, *13*, 348–372.
81. Kajopoulos, J.; Wong, A.H.Y.; Yuen, A.W.C.; Dung, T.A.; Kee, T.Y.; Wykowska, A. Robot-assisted training of joint attention skills in children diagnosed with autism. In Proceedings of the International Conference on Social Robotics, Paris, France, 26–30 October 2015; Springer: Cham, Switzerland, 2015; pp. 296–305.
82. Bharatharaj, J.; Huang, L.; Al-Jumaily, A.M.; Krageloh, C.; Elara, M.R. Experimental evaluation of parrot-inspired robot and adapted model-rival method for teaching children with autism. In Proceedings of the 2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV), Phuket, Thailand, 13–15 November 2016; pp. 1–6.
83. Bharatharaj, J.; Huang, L.; Mohan, R.E.; Al-Jumaily, A.; Krägeloh, C. Robot-Assisted Therapy for Learning and Social Interaction of Children with Autism Spectrum Disorder. *Robotics* **2017**, *6*, 4.
84. Koch, S.A.; Stevens, C.E.; Clesi, C.D.; Lebersfeld, J.B.; Sellers, A.G.; McNew, M.E.; Biasini, F.J.; Amthor, F.R.; Hopkins, M.I. A Feasibility Study Evaluating the Emotionally Expressive Robot SAM. *Int. J. Soc. Robot.* **2017**, *9*, 601–613.
85. Warren, Z.E.; Zheng, Z.; Swanson, A.R.; Bekele, E.; Zhang, L.; Crittendon, J.A.; Weitlauf, A.F.; Sarkar, N. Can robotic interaction improve joint attention skills? *J. Autism Dev. Disord.* **2015**, *45*, 3726.

86. Wiese, E.; Müller, H.J.; Wykowska, A. Using a gaze-cueing paradigm to examine social cognitive mechanisms of individuals with autism observing robot and human faces. In Proceedings of the International Conference on Social Robotics, Sydney, Australia, 27–29 October 2014; Springer: Cham, Switzerland, 2014; pp. 370–379.
87. Hirokawa, M.; Funahashi, A.; Pan, Y.; Itoh, Y.; Suzuki, K. Design of a robotic agent that measures smile and facing behavior of children with Autism Spectrum Disorder. In Proceedings of the 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), New York, NY, USA, 26–31 August 2016; pp. 843–848.
88. Yun, S.S.; Choi, J.; Park, S.K. Robotic behavioral intervention to facilitate eye contact and reading emotions of children with autism spectrum disorders. In Proceedings of the 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), New York, NY, USA, 26–31 August 2016; pp. 694–699.
89. Yun, S.S.; Park, S.K.; Choi, J. A robotic treatment approach to promote social interaction skills for children with autism spectrum disorders. In Proceedings of the 2014 RO-MAN: The 23rd IEEE International Symposium on Robot and Human Interactive Communication, Edinburgh, UK, 25–29 August 2014; pp. 130–134.
90. Hirose, J.; Hirokawa, M.; Suzuki, K. Robotic gaming companion to facilitate social interaction among children. In Proceedings of the 2014 RO-MAN: The 23rd IEEE International Symposium on Robot and Human Interactive Communication, Edinburgh, UK, 25–29 August 2014; pp. 63–68.
91. Mavadati, S.M.; Feng, H.; Gutierrez, A.; Mahoor, M.H. Comparing the gaze responses of children with autism and typically developed individuals in human-robot interaction. In Proceedings of the 2014 14th IEEE-RAS International Conference on Humanoid Robots (Humanoids), Madrid, Spain, 18–20 November 2014; pp. 1128–1133.
92. Anzalone, S.M.; Boucenna, S.; Cohen, D.; Chetouani, M. Autism assessment through a small humanoid robot. In Proceedings of the HRI: A Bridge between Robotics and Neuroscience, Workshop of the 9th ACM/IEEE International Conference on Human-Robot Interaction, Bielefeld, Germany, 3–6 March 2014; pp. 1–2.
93. Anzalone, S.M.; Tilmont, E.; Boucenna, S.; Xavier, J.; Jouen, A.L.; Bodeau, N.; Maharatna, K.; Chetouani, M.; Cohen, D.; Group, M.S.; et al. How children with autism spectrum disorder behave and explore the 4-dimensional (spatial 3D+ time) environment during a joint attention induction task with a robot. *Res. Autism Spectr. Disord.* **2014**, *8*, 814–826.
94. Conti, D.; Di Nuovo, S.; Buono, S.; Trubia, G.; Di Nuovo, A. Use of robotics to stimulate imitation in children with Autism Spectrum Disorder: A pilot study in a clinical setting. In Proceedings of the 2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Kobe, Japan, 31 August–4 September 2015; pp. 1–6.
95. Srinivasan, S.M.; Park, I.K.; Neelly, L.B.; Bhat, A.N. A comparison of the effects of rhythm and robotic interventions on repetitive behaviors and affective states of children with Autism Spectrum Disorder (ASD). *Res. Autism Spectr. Disord.* **2015**, *18*, 51–63.
96. Srinivasan, S.M.; Eigsti, I.M.; Gifford, T.; Bhat, A.N. The effects of embodied rhythm and robotic interventions on the spontaneous and responsive verbal communication skills of children with Autism Spectrum Disorder (ASD): A further outcome of a pilot randomized controlled trial. *Res. Autism Spectr. Disord.* **2016**, *27*, 73–87.
97. Srinivasan, S.M.; Kaur, M.; Park, I.K.; Gifford, T.D.; Marsh, K.L.; Bhat, A.N. The effects of rhythm and robotic interventions on the imitation/praxis, interpersonal synchrony, and motor performance of children with autism spectrum disorder (ASD): A pilot randomized controlled trial. *Autism Res. Treat.* **2015**, *2015*, 736516.
98. Mavadati, S.M.; Feng, H.; Salvador, M.; Silver, S.; Gutierrez, A.; Mahoor, M.H. Robot-based therapeutic protocol for training children with Autism. In Proceedings of the 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), New York, NY, USA, 26–31 August 2016; pp. 855–860.
99. Chevalier, P.; Martin, J.C.; Isableu, B.; Bazile, C.; Iacob, D.O.; Tapus, A. Joint Attention using Human-Robot Interaction: Impact of sensory preferences of children with autism. In Proceedings of the 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), New York, NY, USA, 26–31 August 2016; pp. 849–854.
100. Boucenna, S.; Anzalone, S.; Tilmont, E.; Cohen, D.; Chetouani, M. Learning of social signatures through imitation game between a robot and a human partner. *IEEE Trans. Auton. Ment. Dev.* **2014**, *6*, 213–225.

101. Ranatunga, I.; Torres, N.A.; Patterson, R.; Bugnariu, N.; Stevenson, M.; Popa, D.O. RoDiCA: a human-robot interaction system for treatment of childhood autism spectrum disorders. In Proceedings of the 5th International Conference on Pervasive Technologies Related to Assistive Environments, Heraklion, Greece, 6–8 June 2012; p. 50.
102. Costa, S.C.; Soares, F.O.; Pereira, A.P.; Moreira, F. Constraints in the design of activities focusing on emotion recognition for children with ASD using robotic tools. In Proceedings of the 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechanics, Rome, Italy, 24–27 June 2012; pp. 1884–1889.
103. Salvador, M.J.; Silver, S.; Mahoor, M.H. An emotion recognition comparative study of autistic and typically-developing children using the zeno robot. In Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA), Seattle, WA, USA, 26–30 May 2015; pp. 6128–6133.
104. Chevalier, P.; Martin, J.C.; Isableu, B.; Bazile, C.; Tapus, A. Impact of sensory preferences of individuals with autism on the recognition of emotions expressed by two robots, an avatar, and a human. *Auton. Robot.* **2017**, *41*, 613–635.
105. Khosla, R.; Nguyen, K.; Chu, M.T. Socially Assistive Robot Enabled Home-Based Care for Supporting People with Autism. In Proceedings of the Pacific Asia Conference on Information Systems, Singapore, 5–9 July 2015; p. 12.
106. Dehkordi, P.S.; Moradi, H.; Mahmoudi, M.; Pouretmad, H.R. The design, development, and deployment of RoboParrot for screening autistic children. *Int. J. Soc. Robot.* **2015**, *7*, 513–522.
107. Mazzei, D.; Greco, A.; Lazzeri, N.; Zaraki, A.; Lanata, A.; Igliazzi, R.; Mancini, A.; Stoppa, F.; Scilingo, E.P.; Muratori, F.; et al. Robotic social therapy on children with autism: preliminary evaluation through multi-parametric analysis. In Proceedings of the International Conference on Social Computing, Amsterdam, The Netherlands, 3–5 September 2012; pp. 766–771.
108. Cao, H.L.; Esteban, P.G.; De Beir, A.; Simut, R.; Van De Perre, G.; Vanderborght, B. A platform-independent robot control architecture for multiple therapeutic scenarios. *arXiv* **2016**, arXiv:1607.04971.
109. Cootes, T.F.; Edwards, G.J.; Taylor, C.J. Active appearance models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 681–685.
110. Cazzato, D.; Mazzeo, P.L.; Spagnolo, P.; Distanto, C. Automatic joint attention detection during interaction with a humanoid robot. In Proceedings of the International Conference on Social Robotics, Paris, France, 26–30 October 2015; Springer: Cham, Switzerland, 2015; pp. 124–134.
111. Nunez, E.; Matsuda, S.; Hirokawa, M.; Suzuki, K. Humanoid robot assisted training for facial expressions recognition based on affective feedback. In Proceedings of the International Conference on Social Robotics, Paris, France, 26–30 October 2015; Springer: Cham, Switzerland, 2015; pp. 492–501.
112. Leo, M.; Del Coco, M.; Carcagni, P.; Distanto, C.; Bernava, M.; Pioggia, G.; Palestra, G. Automatic emotion recognition in robot-children interaction for asd treatment. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Santiago, Chile, 7–13 December 2015; pp. 145–153.
113. Ponce, P.; Molina, A.; Grammatikou, D. Design based on fuzzy signal detection theory for a semi-autonomous assisting robot in children autism therapy. *Comput. Hum. Behav.* **2016**, *55*, 28–42.
114. Yun, S.S.; Kim, H.; Choi, J.; Park, S.K. A robot-assisted behavioral intervention system for children with autism spectrum disorders. *Robot. Auton. Syst.* **2016**, *76*, 58–67.
115. Simut, R.; Van de Perre, G.; Costescu, C.; Saldien, J.; Vanderfaellie, J.; David, D.; Lebefer, D.; Vanderborght, B. Probogotchi: A novel edutainment device as a bridge for interaction between a child with asd and the typically developed sibling. *J. Evid.-Based Psychother.* **2016**, *16*, 91.
116. Boccanfuso, L.; Scarborough, S.; Abramson, R.K.; Hall, A.V.; Wright, H.H.; O’Kane, J.M. A low-cost socially assistive robot and robot-assisted intervention for children with autism spectrum disorder: Field trials and lessons learned. *Auton. Robot.* **2017**, *41*, 637–655.
117. Liu, R.; Salisbury, J.P.; Vahabzadeh, A.; Sahin, N.T. Feasibility of an autism-focused augmented reality smartglasses system for social communication and behavioral coaching. *Front. Pediatr.* **2017**, *5*, 145.
118. Barakova, E.I.; Bajracharya, P.; Willemsen, M.; Lourens, T.; Huskens, B. Long-term LEGO therapy with humanoid robot for children with ASD. *Expert Syst.* **2015**, *32*, 698–709.

119. Boccanfuso, L.; Barney, E.; Foster, C.; Ahn, Y.A.; Chawarska, K.; Scassellati, B.; Shic, F. Emotional robot to examine different play patterns and affective responses of children with and without ASD. In Proceedings of the 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, New Zealand, 7–10 March 2016; pp. 19–26.
120. Rudovic, O.; Lee, J.; Mascarell-Maricic, L.; Schuller, B.W.; Picard, R.W. Measuring Engagement in Robot-Assisted Autism Therapy: A Cross-Cultural Study. *Front. Robot. AI* **2017**, *4*, 36.
121. KB, P.R.; Lahiri, U. Design of Eyegaze-sensitive Virtual Reality Based Social Communication Platform for Individuals with Autism. In Proceedings of the 2016 7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS), Bangkok, Thailand, 25–27 January 2016; pp. 301–306.
122. Noris, B.; Nadel, J.; Barker, M.; Hadjikhani, N.; Billard, A. Investigating gaze of children with ASD in naturalistic settings. *PLoS ONE* **2012**, *7*, e44144.
123. Courgeon, M.; Rautureau, G.; Martin, J.C.; Grynszpan, O. Joint attention simulation using eye-tracking and virtual humans. *IEEE Trans. Affect. Comput.* **2014**, *5*, 238–250.
124. Esubalew, T.; Lahiri, U.; Swanson, A.R.; Crittendon, J.A.; Warren, Z.E.; Sarkar, N. A step towards developing adaptive robot-mediated intervention architecture (ARIA) for children with autism. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2013**, *21*, 289–299.
125. Bekele, E.; Crittendon, J.A.; Swanson, A.; Sarkar, N.; Warren, Z.E. Pilot clinical application of an adaptive robotic system for young children with autism. *Autism* **2014**, *18*, 598–608.
126. Caruana, N.; McArthur, G.; Woolgar, A.; Brock, J. Detecting communicative intent in a computerised test of joint attention. *PeerJ* **2017**, *5*, e2899.
127. Grynszpan, O.; Nadel, J.; Martin, J.C.; Simonin, J.; Bailleul, P.; Wang, Y.; Gepner, D.; Le Barillier, F.; Constant, J. Self-monitoring of gaze in high functioning autism. *J. Autism Dev. Disord.* **2012**, *42*, 1642–1650.
128. Zheng, Z.; Fu, Q.; Zhao, H.; Swanson, A.R.; Weitlauf, A.S.; Warren, Z.E.; Sarkar, N. Design of an Autonomous Social Orienting Training System (ASOTS) for Young Children With Autism. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, *25*, 668–678.
129. Zheng, Z.; Nie, G.; Swanson, A.; Weitlauf, A.; Warren, Z.; Sarkar, N. Longitudinal Impact of Autonomous Robot-Mediated Joint Attention Intervention for Young Children with ASD. In Proceedings of the International Conference on Social Robotics, Kansas City, MO, USA, 1–3 November 2016; Springer: Cham, Switzerland, 2016; pp. 581–590.
130. Zhang, L.; Wade, J.; Swanson, A.; Weitlauf, A.; Warren, Z.; Sarkar, N. Cognitive state measurement from eye gaze analysis in an intelligent virtual reality driving system for autism intervention. In Proceedings of the 2015 International Conference on Affective Computing and Intelligent Interaction (ACII), Xi'an, China, 21–24 September 2015; pp. 532–538.
131. Gyori, M.; Borsos, Z.; Stefanik, K.; Csákvári, J. Data Quality as a Bottleneck in Developing a Social-Serious-Game-Based Multi-modal System for Early Screening for High Functioning Cases of Autism Spectrum Condition. In Proceedings of the International Conference on Computers Helping People with Special Needs, Linz, Austria, 13–15 July 2016; Springer: Cham, Switzerland, 2016; pp. 358–366.
132. Lun, R.; Zhao, W. A Survey of Applications and Human Motion Recognition with Microsoft Kinect. *Int. J. Pattern Recognit. Artif. Intell.* **2015**, *29*, 1555008.
133. Zhao, W. A concise tutorial on human motion tracking and recognition with Microsoft Kinect. *Sci. China Inf. Sci.* **2016**, *59*, 93101.
134. Christinaki, E.; Vidakis, N.; Triantafyllidis, G. A novel educational game for teaching emotion identification skills to preschoolers with autism diagnosis. *Comput. Sci. Inf. Syst.* **2014**, *11*, 723–743.
135. Zheng, Z.; Das, S.; Young, E.M.; Swanson, A.; Warren, Z.; Sarkar, N. Autonomous robot-mediated imitation learning for children with autism. In Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China, 31 May–7 June 2014; pp. 2707–2712.
136. Ge, B.; Park, H.W.; Howard, A.M. Identifying Engagement from Joint Kinematics Data for Robot Therapy Prompt Interventions for Children with Autism Spectrum Disorder. In Proceedings of the International Conference on Social Robotics, Kansas City, MO, USA, 1–3 November 2016; Springer: Cham, Switzerland, 2016; pp. 531–540.
137. Yun, S.S.; Choi, J.; Park, S.K.; Bong, G.Y.; Yoo, H. Social skills training for children with autism spectrum disorder using a robotic behavioral intervention system. *Autism Res.* **2017**, *10*, 1306–1323.

138. Uzuegbunam, N.; Wong, W.H.; Cheung, S.c.S.; Ruble, L. MEBook: Kinect-based self-modeling intervention for children with autism. In Proceedings of the 2015 IEEE International Conference on Multimedia and Expo (ICME), Turin, Italy, 29 June–3 July 2015; pp. 1–6.
139. Kuriakose, S.; Kunche, S.; Narendranath, B.; Jain, P.; Sonker, S.; Lahiri, U. A step towards virtual reality based social communication for children with Autism. In Proceedings of the 2013 International Conference on Control, Automation, Robotics and Embedded Systems (CARE), Jabalpur, India, 16–18 December 2013; pp. 1–6.
140. Di Palma, S.; Tonacci, A.; Narzisi, A.; Domenici, C.; Pioggia, G.; Muratori, F.; Billeci, L.; The MICHELANGELO study group. Monitoring of autonomic response to sociocognitive tasks during treatment in children with autism spectrum disorders by wearable technologies: A feasibility study. *Comput. Biol. Med.* **2017**, *85*, 143–152.
141. Zhang, L.; Wade, J.; Bian, D.; Fan, J.; Swanson, A.; Weitlauf, A.; Warren, Z.; Sarkar, N. Cognitive load measurement in a virtual reality-based driving system for autism intervention. *IEEE Trans. Affect. Comput.* **2017**, *8*, 176–189.
142. White, S.W.; Richey, J.A.; Gracanin, D.; Coffman, M.; Elias, R.; LaConte, S.; Ollendick, T.H. Psychosocial and Computer-Assisted Intervention for College Students with Autism Spectrum Disorder: Preliminary Support for Feasibility. *Educ. Train. Autism Dev. Disabil.* **2016**, *51*, 307.
143. Bekele, E.; Wade, J.; Bian, D.; Fan, J.; Swanson, A.; Warren, Z.; Sarkar, N. Multimodal adaptive social interaction in virtual environment (MASI-VR) for children with Autism spectrum disorders (ASD). In Proceedings of the 2016 IEEE Virtual Reality (VR), Greenville, SC, USA, 19–23 March 2016; pp. 121–130.
144. Fan, J.; Wade, J.W.; Bian, D.; Key, A.P.; Warren, Z.E.; Mion, L.C.; Sarkar, N. A Step towards EEG-based brain computer interface for autism intervention. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 3767–3770.
145. Özcan, B.; Caligiore, D.; Sperati, V.; Moretta, T.; Baldassarre, G. Transitional wearable companions: A novel concept of soft interactive social robots to improve social skills in children with autism spectrum disorder. *Int. J. Soc. Robot.* **2016**, *8*, 471–481.
146. Bian, D.; Wade, J.; Warren, Z.; Sarkar, N. Online Engagement Detection and Task Adaptation in a Virtual Reality Based Driving Simulator for Autism Intervention. In Proceedings of the International Conference on Universal Access in Human-Computer Interaction, Toronto, ON, Canada, 17–22 July 2016; Springer: Cham, Switzerland, 2016; pp. 538–547.
147. Bakeman, R. Behavioral observation and coding. In *Handbook of Research Methods in Social and Personality Psychology*; Cambridge University Press: New York, NY, USA, 2000; pp. 138–159.
148. Weick, K.E. Systematic observational methods. In *The Handbook of Social Psychology*; John Wiley and Sons: Hoboken, NJ, USA, 1968; Volume 2, pp. 357–451.
149. Khosla, R.; Nguyen, K.; Chu, M.T. Service personalisation of assistive robot for autism care. In Proceedings of the IECON 2015-41st Annual Conference of the IEEE Industrial Electronics Society, Yokohama, Japan, 9–12 November 2015; pp. 002088–002093.
150. Mourning, R.; Tang, Y. Virtual reality social training for adolescents with high-functioning autism. In Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary, 9–12 October 2016; pp. 004848–004853.
151. Bekele, E.; Zheng, Z.; Swanson, A.; Crittendon, J.; Warren, Z.; Sarkar, N. Understanding how adolescents with autism respond to facial expressions in virtual reality environments. *IEEE Trans. Vis. Comput. Graph.* **2013**, *19*, 711–720.
152. Kaboski, J.R.; Diehl, J.J.; Beriont, J.; Crowell, C.R.; Villano, M.; Wier, K.; Tang, K. Brief report: A pilot summer robotics camp to reduce social anxiety and improve social/vocational skills in adolescents with ASD. *J. Autism Dev. Disord.* **2015**, *45*, 3862–3869.
153. Newbutt, N.; Sung, C.; Kuo, H.J.; Leahy, M.J. The potential of virtual reality technologies to support people with an autism condition: A case study of acceptance, presence and negative effects. *Ann. Rev. Cyber Ther. Telemed. (ARCTT)* **2016**, *14*, 149–154.
154. Ekman, P.; Friesen, W.V. *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues*; ISHK: Los Altos, CA, USA, 2003.
155. Benton, A.L. The neuropsychology of facial recognition. *Am. Psychol.* **1980**, *35*, 176.

156. Gresham, F.M.; Elliott, S.N. *Social Skills Rating System: Manual*; American Guidance Service: Circle Pines, MN, USA, 1990.
157. Suskind, R.; Nguyen, J.; Patterson, S.; Springer, S.; Fanty, M. Guided Personal Companion. U.S. Patent 9,710,613, 18 July 2017.
158. Liu, X.; Zhao, W. Buddy: A Virtual Life Coaching System for Children and Adolescents with High Functioning Autism. In Proceedings of the IEEE Cyber Science and Technology Congress, Orlando, FL, USA, 5–9 November 2017.
159. Bernardini, S.; Porayska-Pomsta, K.; Smith, T.J. ECHOES: An intelligent serious game for fostering social communication in children with autism. *Inf. Sci.* **2014**, *264*, 41–60.
160. Esteban, P.G.; Baxter, P.; Belpaeme, T.; Billing, E.; Cai, H.; Cao, H.L.; Coeckelbergh, M.; Costescu, C.; David, D.; De Beir, A.; et al. How to build a supervised autonomous system for robot-enhanced therapy for children with autism spectrum disorder. *Paladyn J. Behav. Robot.* **2017**, *8*, 18–38.
161. Luo, X.; Deng, J.; Liu, J.; Wang, W.; Ban, X.; Wang, J.H. A quantized kernel least mean square scheme with entropy-guided learning for intelligent data analysis. *China Commun.* **2017**, *14*, 1–10.
162. Luo, X.; Deng, J.; Wang, W.; Wang, J.H.; Zhao, W. A Quantized Kernel Learning Algorithm Using a Minimum Kernel Risk-Sensitive Loss Criterion and Bilateral Gradient Technique. *Entropy* **2017**, *19*, 365.
163. Luo, X.; Lv, Y.; Zhou, M.; Wang, W.; Zhao, W. A laguerre neural network-based ADP learning scheme with its application to tracking control in the Internet of Things. *Pers. Ubiquitous Comput.* **2016**, *20*, 361–372.
164. Luo, X.; Xu, Y.; Wang, W.; Yuan, M.; Ban, X.; Zhu, Y.; Zhao, W. Towards Enhancing Stacked Extreme Learning Machine With Sparse Autoencoder by Correntropy. *J. Frankl. Inst.* **2017**, doi:10.1016/j.jfranklin.2017.08.014.
165. Luo, X.; Zhang, D.; Yang, L.T.; Liu, J.; Chang, X.; Ning, H. A kernel machine-based secure data sensing and fusion scheme in wireless sensor networks for the cyber-physical systems. *Future Gener. Comput. Syst.* **2016**, *61*, 85–96.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).