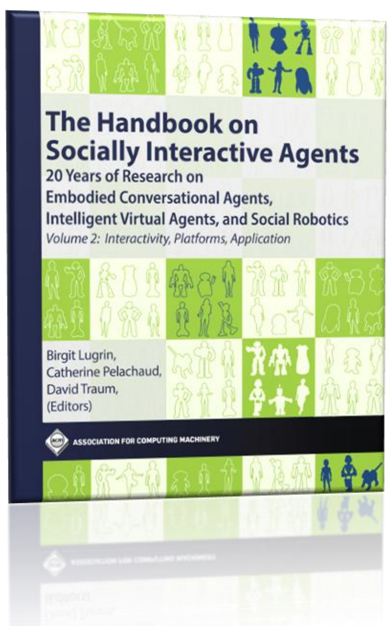
Autism and Socially Interactive Agents

Jacqueline Nadel, Ouriel Grynszpan, and Jean-Claude Martin



Author note:

This is a preprint. The final article is published in "The Handbook on Socially Interactive Agents" by ACM.

Citation information:

J. Nadel, O. Grynszpan, and J.-C. Martin (2022). Autism and socially interactive agents . In B. Lugrin, C. Pelachaud, D. Traum (Eds.), *The Handbook on Socially Interactive Agents – 20 Years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics*, Volume 2: Interactivity, Platforms, Application (pp. 437-562). ACM.

DOI of the final chapter: 10.1145/3563659.3563673

DOI of volume 2 of the handbook: 10.1145/3563659

Correspondence concerning this article should be addressed to Jacqueline Nadel (jacqueline.nadel06@gmail.com), Ouriel Grynszpan (ouriel.grynszpan@u-psud.fr), and Jean-Claude Martin (martin@limsi.fr)

25 Autism and Socially Interactive Agents

Jacqueline Nadel, Ouriel Grynszpan, Jean-Claude Martin

25.1 Motivation

Autism Spectrum Disorder (ASD) is defined in the Diagnostic and Statistical Manual of Mental Disorders of the American Psychiatric Association in the section on neurodevelopmental disorders. Among other things, individuals with ASD may exhibit "persistent deficits in social communication and social interaction in multiple contexts, including deficits in social reciprocity, in non-verbal communicative behaviors used in social interaction, and in skills in developing, maintaining, and understanding interpersonal relationships" [APA, 2013]. Besides social impairment, individuals with ASD exhibit motor problems that impede normal interactions with the physical world and generate poor autonomy.

Advantages of Socially Interactive Agents

Socially Interactive Agents (SIA) display verbal and non-verbal communicative behaviors that can be controlled while offering a wide range of possible degrees of realism. They induce socially relevant cognitive processes in humans who interact with them and are therefore very valuable as investigation tools to study ASD [Grynszpan et al., 2009; Recht & Grynszpan, 2019; Schilbach et al., 2010]. They are especially interesting to analyze the live dynamics of social interactions, which are associated with some of the most profound and persistent impairments in ASD [Senju et al., 2009].

Deficits in social communication and social interaction

While many tablet applications aim to train social skills in people with ASD, most of these applications currently available to the general public do not involve real social interaction: the majority of these apps do not analyze user's behavior during an interaction, since the user must make choices from menus or point to graphical elements on the tablet. A few apps help monitor user's behavior and emotional states. Furthermore, they do not allow the collection of experimentally controlled measures aimed at better understanding these disorders.

Beyond the numerous applications available on tablets for social skills training, several reviews of the state-of-the-art and meta-analyses have already been published on research into new technologies applied to autism. A review has been published on the training of skills assisted by digital technologies in autism [Grynszpan et al., 2014]. This review shows that different social skills are generally targeted (such as the recognition of facial or bodily expressions of emotions) and that different technologies are used (tablets, robotics, collaborative platforms, eye tracking). The authors report that there is evidence in favor of the effectiveness of these trainings but that nevertheless, questions remain regarding the heterogeneity of the methods used, the impact of human caregivers, the maintenance of the effects and their generalization to everyday life skills. Finally, risks of excessive use of IT tools and potential isolation are underlined.

In their review of 40 technological applications to train children with autism in social skills, DiGennaro Reed and colleagues mention the following social skills in descending order of use: initiating conversation, play, social conventions, responding to others, non-verbal behaviors, social problem solving, regulating emotions and reciprocity, and friendship relationships [DiGennaro Reed et al., 2011]. Obviously, all these technological applications concern verbal children with ASD, as nonverbal behaviors are not among the more frequently used. Most of the studies reported in this review used videos, took place in schools, and targeted several social behaviors together. Serious games with visual support aim to train cognitive, social and sensory capacities in a playful way. GOLIAH is a serious game delivered at home which combines a pedagogical program and a e-therapy of joint attention and imitation. A pilot study suggests the feasibility of using the developed gaming platform for home-based intensive intervention. However, the overall capability of the platform in delivering intervention needs to be assessed in a bigger open trial [Bono et al., 2016].

Other reviews of the state of the art were dedicated to specific types of technology. For instance, a systematic review of about thirty serious games to teach people with ASD to manage social interactions was published [Grossard et al., 2017]. The serious games reviewed seem promising because they provide a wide variety of skills that can be practiced in a variety of real-life situations. However, they also have limitations: most of them were designed for people with high functioning autism, their clinical validation did not meet the standards of evidence-based medicine, the design stage of the game engine was not described, and, unfortunately, clinical validation and playability were usually not given equal consideration.

Motor coordination deficit

A motor coordination deficit is now well documented as a cardinal feature of ASD across age, gender and severity of syndrome [Fournier et al., 2010]. Motor coordination disorder impedes physical, social and cognitive capacities. Embedded and embodied as it is, virtual reality (VR) offers efficient ways to remediate to basic motor deficits via training. Notably VR can benefit to nonverbal people with limited representational abilities thanks to the use of graspable objects that add to the feeling of realistic presence. Cameras of Kinect type allow to train gestures and bodily recognition using a gesture interactive game-based learning approach to improve preschool children's learning performance and motor skills [Hsiao & Chen, 2016]. Children with neurodevelopmental disorder are proposed this kind of game with the aim to evaluate motor and cognitive performance [Kourakli et al., 2017]. Specialists of Kinect games propose minigames that evolve according to the children's performance [Bartoli et al., 2014]. Within the same framework but more fundamental and developmental in option, Pictogram Room was designed to develop motor and social skills in children with ASD based on the training of imitation, joint attention and body knowledge. As in the games mentioned above, there are no humanoid virtual agents, but instead virtual avatars of the child and of the adult. Virtual objects moving on the screen allow to train joint attention and body performance [Herrera et al., 2012]. A calibration of the avatar as a monochrome exoskeleton in wire is registered at the beginning of the session. However, although minimalist, this representation of self as an exoskeleton is not understandable to all low-functioning children with ASD. Indeed, such representation requires meta self-recognition (i.e. you recognize yourself in a symbolic character that does not resemble to you but acts like you). To help achieve cognitive self-recognition, an intervention is proposed using Kinect to alternately offer three representations of self to the child: a real image, a “silhouette” and an exoskeleton [Nadel & Poli, 2018].

25.2 Models and Approaches

New technologies offer a revival in all disciplines. They are involved in the adaptation and creation of physiological, cerebral and behavioral exploration techniques. Applications, architectures, designs are based on a scientific body: the sciences of the artificial, combining computer science, cybernetics, electronics, artificial intelligence, cognitive sciences, and many other disciplines.

Animated characters are often used to design animated videos that will then be selected and displayed during the interaction, or even react in real time to the user's behaviors during the interaction. Some of these prototypes have been designed for and evaluated by people with ASD. Other prototypes have been developed for neurotypical users only and then adapted for autistic users. Depending on the computer platforms and research prototypes used, these animated characters can be more or less expressive, and more or less interactive in real time. Several researchers focus on animated characters because they are particularly relevant for the simulation and training of social skills. Indeed, they have multiple advantages over pre-recorded videos of real humans. Their non-verbal expressions (e.g. facial expressions, gaze, postures) can be controlled more finely. When they are truly interactive (and therefore not just pre-recorded videos), they can be used to train continuous and finely tuned behaviors (e.g. eye tracking as we will see later). They are thus seen as a good compromise between experimental control and ecological validity to study and simulate social interactions using verbal and non-verbal behaviors.

The interest in imitation as a form of learning that could be endowed with artificial systems developed as early as the second half of the 20th century. By putting imitation learning techniques in computers, the door was open for many applications allowing the automatic acquisition of capabilities and behaviors on the basis of programs via human-computer, human-robot and robot-robot interfaces where new robots can acquire know-how on the basis of reinforcement by observing the behavior of other robots or humans. To get out of this vision of engineers, several theories stand up, among them cybernetics, the theory of dynamic systems and its cognitive component: enaction. They called for the introduction of the notion of systems located and embodied. An interdisciplinary constructivist option was born, seeking to understand the mechanisms that can generate a phenomenon [Nehaniv & Dautenhahn, 2007]. Questions ensue such as: can an explanation proposed by a theory be validated by constructing an artifact that exhibits the behavior suggested by the theory? Is the theory or model detailed enough to build artificial systems that embody it? Are there multiple non-equivalent ways to build achievements of these mechanisms? Are there gaps in the construction of certain models? Does the model lead to predictions about the behavior of organisms? This last point is of interest in the realm of ASD. Indeed autistic behaviors are often peculiar and general models may be irrelevant to make valid predictions about how people with ASD react to what. Specific models are thus of paramount importance: Knowing for instance that people with ASD, even low-functioning ones, recognize be imitated and respond positively to this social initiation, and that they can imitate spontaneously gestures and actions that are in their motor repertory [Nadel, 2015], imitation becomes an interesting model of social development in ASD [Nadel, 2014] which can inspire the design of SIA based interactive systems.

Animated agents and other types of interactive devices are more and more often used in social neuroscience studies in association with various measures and devices: fMRI, EEG, ECG, ... [Georgescu et al., 2014]. Such studies aim at answering research questions such as “How does the brain allow us to interact with others?” A good example is the series of studies designed by Kelso’s team. They combine the human dynamic clamp, a novel paradigm for studying realistic social behavior, with high-resolution electroencephalography (Dumas et al, 2019 ; Kelso et al.,2009 ; Tognoli & Kelso, 2015]. In one of the studies, participants were asked to interact via a finger device with a partner ‘s finger, without knowing whether the partner was a human or a virtual partner (VP). The VP was randomly assigned a cooperative or competitive behavior for two halves of each trial, giving four pseudo-randomized types of trials: cooperation throughout, competition throughout, switch from cooperation to competition, and finally switch from competition to cooperation. After the interaction periods, the participants were asked to judge the humanness of the VP through a binary choice (0: machine and 1: human). At the behavioral level, results show a link between the attribution of humanness to the VP and cooperative phases. At the brain level, judgment of humanness and cooperation of others modulate the functional connectivity between areas and reveal how distributed neural dynamics integrates information from “low-level” sensorimotor mechanisms and “high-level” social cognition to support the realistic social behaviors that play out in real time during interactive scenarios [Dumas et al., 2019]. The protocol is currently used with children with ASD to address their capacity to attribute humanness to a partner that they do not see, based on joint synchronized action of the partner. Knowing that brains synchronize during a synchronic action between two persons [Dumas et al., 2010], the protocol tests how far a virtual thumb is enough to test social attribution at the behavioral and brain levels.

25.3 History / Overview

25.3.1 Robots

Following the automatons that multiplied from the 17th century onwards, robots appeared in the 20th century. Robots clearly differ from the automatons by their autonomous feature. Their appearance is linked to the development of cybernetics, founded by Norbert Wiener (1948) and applied notably by William Grey Walter (1950) with his famous cybertortues, artificially reproducing basic Pavlov's conditioned reflexes. From animal simulation, it turns to human simulation with the development of anthropoid robots. Cybernetics, 'science of controlled

analogies between organisms and machines', and its extension, neurocybernetics, evolved from the 1990s by coupling with artificial intelligence.
A good illustration of this coupling is given by the development of *Affective computing*, defined as the study and development of systems and devices that can recognize, interpret, process, and simulate human affects. It is an interdisciplinary field spanning computer science, psychology, and cognitive science, first framed by the manifesto for affective learning led by Rosalind Picard [Picard et al., 2004]. Within this framework, the affective robot Jibo was designed in 2014 by Cynthia Breazeal, leader of the Personal Robots Group of the MIT Media Lab. The team is pioneer in the domain of social robotics and communication between robots and persons [Breazeal et al., 2009] and has demonstrated the possible development of affective links between robots and individuals or groups of individuals, in particular families. Now then robots are not only autonomous, they are also intelligent, and even for some of them, intelligent connected robots.

Several studies have shown the potentials of Social Robots (SR)', especially humanoids, to study the abilities of children with ASD to imitate expressions or movements [Billard et al., 2007], which play a key role in social interactions [Nadel, 2014], as we explained above. Dautenhahn's team [Robins & Dautenhahn, 2018] sees its little Robot Kaspar as an aid to parents and teachers. The team was able to show by case studies involving 170 children, how Kaspar (Figure 1) helps to get out of isolation children with severe autism, and encourages them to interact through a variety of abilities such as imitation, smiling contact, joint attention. Interaction was controlled on wireless keyboard by adults or children themselves. The authors highlight the positive findings of an independent team of 54 practitioners on the effects of Kaspar on development [Huijnen et al., 2016]. Though there was no long-term evaluation, at least the many follow-ups demonstrated a notable effect of Kaspar on the frequency and relevance of social behaviors. Moreover, a systematic user-centered perspective was developed.

Figure 1. Kaspar the little social robot designed by Kerstin Dautenhahn. Image © 2021 Jacqueline Nadel.

NAO (Figure 2) has been used in multiple studies with ASD users such as the ASK NAO initiative for educational or intervention tool. As an intervention tool, Nao showed its potentiality to improve stereotyping in children with ASD [Shamsuddin et al., 2012]. Many interventions including Nao, however, are still pilot ones. As an educational tool, Nao has great potential. Its audio speakers and microphones, coupled with algorithms of recognition and verbal production, allow it to locate the origin of a sound and to understand and speak with the appropriate prosody. The educational programs provided by ASK NAO are customizable according to the profile and needs of each child. but the *a la carte* platform requires longitudinal follow-ups. Moreover, its use is not simple and requires a trained staff.

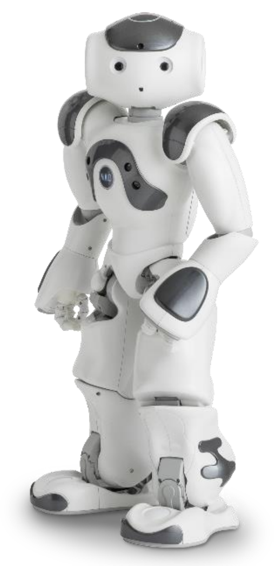


Figure 2. NAO by SoftBank Robotics, an intelligent robot. Image © 2021 SoftBank Robotics.

More generally, a growing pedagogical approach named educational robotics considers robots as *serious toys* to engage students with or without ASD in scientific digital reasoning. An innovating use of educational robotics has included Nao as a silly little boy programmed to make big errors. The students will have to teach Nao the good answers…and it works! [Masson et al., 2017].

A review of approximately 40 studies in the field of social robotics applied to autism between 2006 and 2016 [Pennisi et al., 2016] found positive impacts of robots. For example, some people with ASDs showed better results and showed more social behaviors with a robot than with a person. However, the authors raise the question of the impact of inter-individual differences (gender, IQ, age) and the observability of these behaviors outside of experimental and clinical contexts. In a recent study in social robotics [Yun et al., 2017], the authors trained what they call "basic" skills: eye contact and recognition of facial expressions of emotion. Two conditions were compared (human coach vs. robot coach). The authors observed improvements in both conditions in terms of emotion recognition and eye contact, suggesting that robots may indeed be useful for learning social behaviors and decreasing behavioral and emotional symptoms (since they achieve similar results than those obtained with a human coach).

25.3.2. Virtual Agents

The history of virtual agents starts with artificial intelligence and is totally involved in designing artificial creatures demonstrating social and physical behaviors as close as possible to human behaviors. In one of the earliest work using intelligent virtual humans to train pupils with ASD, Massaro and Bosseler (2006) designed and tested a training method to teach vocabulary with a virtual human head simulating realistic word pronunciation. Mitchell et al. (2007) trained adolescents with ASD to behave in a socially adequate way when looking for a place to sit in a café populated with intelligent virtual agents. Participants could engage in conversation with virtual customers in this virtual café using fixed sets of pre-defined sentences. Researchers reported improvements in social reasoning after training three adolescents with ASD. Grynszpan et al. (2012) investigated social gaze in ASD by providing real-time feedback on their gaze direction to participants who were being addressed by virtual animated human characters. Kandalaft et al. (2013) used an online collaborative virtual environment to simulate social settings (e.g. job interview, conversation with a roommate, negotiating with a salesman) involving the avatars of both clinicians and individuals with ASD. Significant increases on social cognitive measures of theory of mind and emotion recognition, as well as in real life social and occupational functioning were found post-training. These findings suggest that the virtual reality platform is a promising tool for improving social skills, cognition, and functioning in autism.

Simulating realistic social exchanges between users with ASD and virtual agents requires to endow agents with abilities to respond in a socially meaningful way. Bernardini et al. (2014) designed an autonomous virtual agent to train turn-tacking and joint attention in children with ASD. The agent appeared on a large touch-screen area equipped with cameras that monitored the children's gestures. The system could recognize pointing gestures and gaze orientation. The virtual agent's behavior was controlled by an artificial intelligence engine that included an emotional model and a model of the child user. The virtual agent's ability to interact was nevertheless restricted due to the current limitations of artificial intelligence in appropriately modelling social cognitive behavior. To bypass the lack of adequate intelligent model of social behavior, Tartaro (2007) created a virtual character that could be controlled by an experimenter to play with a child with ASD. Kozima et al. (2007) opted for the same strategy with a robot that was remotely controlled by an experimenter to interact with children.

25.3.3. Overview of the use of SIAs in different interactive settings

The aim of research about Socially Interactive Agents (SIA) and Autism is twofold: 1) to allow advanced experimental studies aimed at better understanding some specific deficits of individuals with ASDs (often without a rich situational context), and 2) at a more macroscopic level, to design and evaluate virtual as well as robotic training systems for social skills (often in one or more specific interaction situations). In this section, we review articles describing these two approaches (state-of-the-art experimental studies and training system evaluation).

We propose a reading grid of these different works based on the different types of SIA systems and the different configurations of social interactions they allow. The design of the review that we propose in this paper thus follows Figure 3 (based on the interactive situations with virtual agents [Rist et al., 2003]): (A) non-interactive presentation with a single user and a single SIA (either a virtual agent or a robot), (B) interaction between a user and a SIA, (C) presentation to a user of interactions between several SIAs, (D) interaction between a user and several SIAs that also interact with each other, and (E) interactions between several co-present users interacting with several SIAs that also interact with each other.

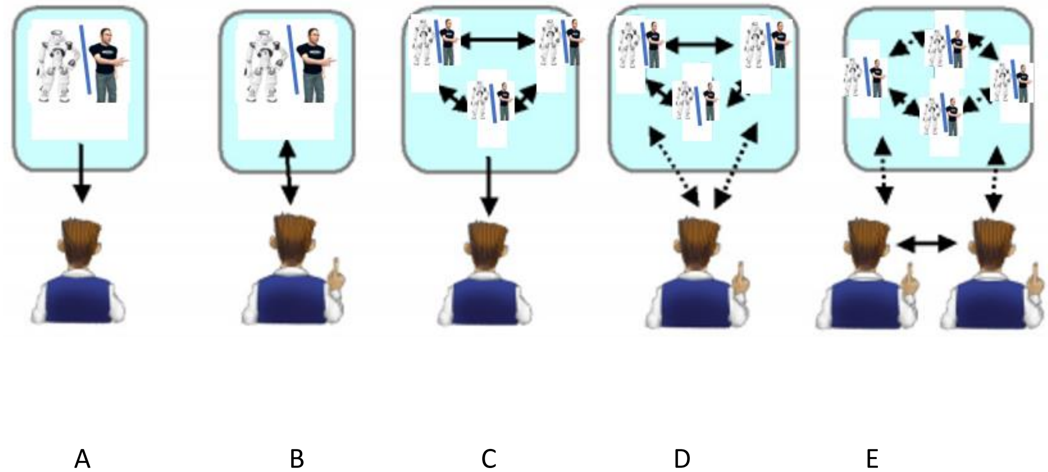


Figure 3: A schematic view of the different types of SIA-based systems (based on (Rist et al. 2003)).

The systems and studies we review are very diverse and show the potential of intelligent virtual and robotic agents for social skills training. Some systems aim at testing the abilities of users with ASD, others at training them. Some systems focus on one modality (e.g. gaze), other systems aim at a combined use of several social skills and several modalities in complex situations (e.g. public speaking).

Non-interactive presentation with a virtual agent and a user

Many studies and applications use pre-calculated videos of virtual (i.e., non-interactive) character animations by presenting them during interaction to people with ASDs. Some serious games involve recognizing facial expressions of emotions located in a graphic environment. JEStiMulE (Educational Game for Multisensory Stimulation of Children with Autism Spectrum Disorder) is a serious game training emotion recognition in social contexts [Serret et al., 2014]. As a multi-sensory tool combining visual, auditory and tactile feedback, it is used as a complement to children's psychotherapy. Children must learn to recognize and anticipate emotions through two phases: 1) recognizing facial expressions (the child must learn to recognize emotions through color coding and tactile patterns sent by a joystick), and 2) associating emotions with context (the child must recognize and anticipate emotions in context).

It should be noted that unlike approaches using the technique known as video modeling, which relies on using videos already recorded to illustrate social behaviors to be learned (with real people who are filmed), these games and applications use animated characters whose non-verbal behavior is more easily controllable and whose more visible expressions may be more relevant for learning. However, these games provide exposure to social behavior and not social interaction per se, since there is no real interaction between the user and the animated characters. The behavior of virtual agents does not change according to the user's actions during the interaction.

There is just a selection of the video/animation that are played at a given moment. The dynamics of the dyadic interaction between the user and the virtual character, when present, stays at a very macroscopic level, which is not the case for the social interactions we have every day with our peers.

Virtual agents expressing basic emotions with their facial expressions, verbalizations and gestures have been used in many experiments with little or no interaction. This is the case of a study on approach and avoidance behaviors in which 19 children with high functioning autism had to identify emotions expressed by a virtual agent and position themselves in relation to the virtual agent using a joystick [Kim et al., 2015]. Using this device, it was observed that, compared to a group of typical children, children with ASD showed less approach behavior towards virtual agents expressing positive emotions. However, there was no difference between children with ASD and typical children in terms of aversive behaviors in relation to virtual agents expressing negative emotions. In this type of setting, consideration of user behaviors is limited to the movement of a joystick. Virtual agents are displayed one after the other and differ only in their appearance and the emotions they display. This experimentation aimed at better understanding approach and avoidance behaviors and was not aimed at learning social skills.

Many studies have used eye-tracking devices to study how people with autism explore a video showing social behaviors. In a more original way, [Grynszpan et al. 2012; Grynszpan & Nadel, 2015] designed and evaluated a system based on gaze-contingent display that allowed real-time blurring of video except for an area centered on the point of the screen gazed at by the user. The authors observed that the more participants used this system to actively view virtual agent facial expressions, the more they were able to describe the virtual agent's expressions in terms of cognitive verbs. Nevertheless, individuals with autism showed lower adaptation to this system of gaze contingent display. Although the device impacted the visibility and display of the animation, it did not change the interactive behavior of the virtual agent in real time.

Interaction between a virtual agent and a user

The studies described in the previous section do not allow the learning of social interaction skills per se since the user's behaviors do not impact the virtual agent's behaviors. These real-time dynamic adjustment skills during social interactions nevertheless need to be trained in people with ASDs. In this section, we describe research prototypes in which the user's behavior impacts the social behaviors of the virtual agent during the interaction.

In learning situations, interaction with peers has shown positive results [Bowman-Perrott et al., 2013]. A system using this track allows a child with autism to play with physical objects such as the components of a child's home, while interacting with a virtual child about the toy [Tartaro, 2007]. Even if physical objects were not interactive, they allowed social interactions to be grounded in the physical and real world.

Interviewing for a job is an example of a task that is recognized as requiring complex social and non-verbal communication skills and eliciting social stress. It is therefore a complex task, particularly for people with high functioning autism. Several research prototypes have been developed to train young adults for job interviews, targeting primarily typical participants. The MACH system [Hoque et al., 2013] controls the behavior of an animated virtual recruiter with the MARC expressive virtual agent platform [Courgeon et al., 2014]. User behaviors are analyzed in terms of smile, vocabulary richness and head movements. The virtual agent displays non-verbal back-channel behaviors for head movements and smiles when the user smiles as well. A study showed that training with this system had a positive impact on the evaluation of candidates' social skills by external assessors. Another system, TARDIS, focused on training to conduct job interviews with virtual recruiters displaying different attitudes (friendly/aggressive, dominant/submissive) [Ben Youssef et al., 2015]. These different systems, although they have received interest from people with ASDs, were not designed specifically for users with ASDs, which limits their potential for use with these users.

However, some job interview training systems have been successfully applied to participants with different pathologies showing problems related to social interactions that make it difficult for them to conduct job interviews. The VR-JIT (Virtual Reality Job Interview Training) system does not use virtual agents but videos of a human actress playing the role of a recruiter. Different elements of the graphical interface are used to help the candidate interpret the behavior of the recruiter or to listen to her own answers again. This system has been evaluated with participants with different pathologies: post-traumatic stress disorder [Smith et al., 2015] or schizophrenia [Smith et al., 2015]. It has also been successfully applied to adults with ASDs [Smith et al., 2014]. Participants using the system showed more improvement than a group of adults with ASDs receiving conventional job interview documentation. The system was also well perceived by users in terms of ease and playfulness. Unfortunately, the interview training systems described above only rarely allow for a continuous dynamic of the virtual agent's behavior based on fine user behaviors such as eye movement, which is nevertheless the case during real social interactions.

Indeed, virtual agents can help investigate social gaze in autism. Impaired ability to use eye communication is a hallmark of autism [APA, 2013]. Gaze based interaction is distinctively human [Kobayashi & Kohshima, 1997] and provides the basis of what is known as joint attention [Emery, 2000]. Joint attention is an umbrella term for behaviors that allow two (or more) human beings to focus their attention on the same object of interest. Impairments in joint attention are among symptoms that are the most specific of autism in toddlers [Dawson et al., 2004]. Several research projects seek to address those impairments. Trepagnier et al. (2006) designed a gaze-contingent virtual agent whose behavior depended on the amount of eye contact with the user. Dratsch et al. (2013) employed an intelligent virtual character to evaluate the ability of people with ASD to detect when someone else’s gaze is directed at them.

The VIGART system analyzed the gaze of participants with autism when they listened to a virtual agent telling a story and displaying facial expressions. This system provided feedback to the user by, for example, encouraging the user to look more at the virtual agent while the virtual agent told the story [Lahiri et al., 2011]. Virtual agents could be customized to look like people the user knew. However, the user's eye behavior had no impact on the virtual agent's behavior and was only used for analysis and for displaying recommendations. The different virtual agents were again displayed sequentially, without any social interaction between them.

Behavioral researchers have expressed interest in gaze-contingent virtual agents to explore the ASD syndrome. Such virtual agents were instrumental to assess the ability of people with ASD in using gaze for deictic communication [Caruana et al., 2017]. Little et al. (2016) relied on gaze-contingent virtual agents to investigate the effect of joint attention on the memory of objects that were gazed at. The MARC interactive virtual character platform combined with eye-tracking devices was used in a series of studies on joint attention to create virtual characters that could continuously follow the user’s gaze in real time [Courgeon et al., 2014], It was used to show that typical people tend to prefer virtual agents that follow their gaze over those that do not [Grynszpan et al., 2017]. A subsequent study based on the same platform revealed that young adults with ASD were less prone than their typical peers to realize that they were leading the gaze of a social partner during joint attention episodes. This line of research shows how using virtual agents can help improving scientific knowledge about ASD.

Systems using multiple virtual agents

Few systems manage multiple virtual agents at the same time when interacting with one or more users (Figure 1). Many systems allow conversations between several virtual agents displayed on the screen at the same time without allowing the user to interact with these agents (Figure 1C). In some cases, these animated presentations result from computer simulations of communication behaviors between virtual agents displayed on the screen (e.g., speech turns and attitudes) [André, 2002]. In particular, these systems have been used to present two different points of view on a commercial product or a football match taking place on the screen.

The sophistication of considering the interactions between several virtual agents and a user varies from system to system. Some systems are limited to allowing a user to interact sequentially with different virtual agents without the system simulating and displaying interactions between virtual agents (or in a very limited way). The aim is then to improve social skills by interactively training in several social situations sequentially, each time with a different virtual character. Thus, the Virtual Reality Social Cognition Training (VR-SCT) system simulates interactions with virtual characters in different situations of the Second Life environment: social introduction and interaction with a friend with common interests, initiating a conversation with a roommate, meeting strangers or friends, negotiating with a salesperson, going to a job interview, dealing with conflicts with co-workers, partying with a friend, consoling a friend, interacting with someone without common interests [Kandalaft et al.,2013]. Eight young adults with high functioning autism used the system for 10 sessions over 5 weeks. The authors observed improvements on measures related to theory of mind, emotion recognition, and in everyday life. A virtual coach is also present in the environment and aims to give immediate feedback during training. Even if this type of system does not allow to simulate and animate very sophisticated interactions between the different virtual agents displayed on the user interface, they still allow to train the person with ASD to interact with different characters who have different appearances and different roles. Some systems display several virtual agents in the same graphical scene (without necessarily simulating social interactions between these virtual agents). Other systems simply display the different virtual agents in different windows. This is the case of the Social Tutor system [Milne et al., 2013] designed and used by children with ASD. Social Tutor involves 3 virtual agents (head and upper body) acting as a virtual teacher, a virtual child with good communicative and social skills, and a virtual child *without* good communicative and social skills, respectively. This is the strategy of learning with a virtual peer that we mentioned earlier. This system aims to combine the learning of social skills with the learning of language skills (greetings, conversations, listening behaviors, and turn management). The three characters are displayed at the same time on the screen but in separate windows and without simulated social interactions between them.

The use of several virtual agents for pedagogy has been exploited in several studies with typical adolescents. For example, it has been observed that having two virtual agents displayed at the same time in the same window and embodying two different roles (expert vs. motivator) produced better results than having a single virtual agent taking on both roles [Baylor & Ebbers, 2003]. However, the system did not simulate interactions between the two virtual agents co-present on the screen. Other systems allow for public speaking training using multiple reactive virtual agents. Public speaking, like job interviews, is a task that requires the mastery of complex and integrated social and communicative skills. For example, the CICERO system analyzes user behavior and uses the results of these analyses to monitor the individual behaviors of several virtual agents simultaneously displayed on a large screen and acting as the audience attending the presentation [Chollet et al., 2015]. Thus, the body postures and gaze directions of these virtual agents can be dynamically changed showing their level of interest in the presentation made by the user according to the quality of the presentation, which is automatically estimated.

Peter Mundy and colleagues used such a virtual environment to study the social attention abilities of 37 children with ASDs and 54 typical children [Jarrold et al, 2013]. The system displayed 9 autonomous virtual agents sitting at a virtual table. The user's attention to the 9 agents while answering questions was analyzed. During this talk, children with ASD watched the agents less than children without ASD. In addition, social attention in the group with ASD was moderated as follows: children with lower IQ, higher social anxiety symptoms, and higher attention disorders showed more atypical social attention.

Even if in these different automatic systems with several virtual agents are displayed at the same time in the graphical scene, there is no simulation or animation of the social interactions between these virtual agents. Moreover, the interaction between the user and the virtual agents remains limited: the user cannot, for example, start a conversation with one of the virtual agents. In order to allow richer interactions, some systems involve an experimenter controlling the system which is no longer completely automatic.

An immersive semi-automatic virtual reality system exploiting several virtual agents expressing emotions in two situations (birthday party and school class) involving 10 social situations was used in conjunction with an automatic analysis of the facial expressions of autistic children

[Lorenzo et al., 2016]. Expressions of emotions detected on the user's face (anger, joy, sadness and surprise) are used to update the virtual social situation and are evaluated in terms of relevance to the situation. The system is not completely automatic: an experimenter is responsible for triggering the display of facial expressions of emotions on the virtual agents according to the emotions expressed by the user. Several virtual agents are present but the system does not explicitly support interactions between agents: it is up to the experimenter to do so if he wishes.

Some systems have focused on the use of speech recognition when interacting with intelligent virtual agents. This is the case of Ada and Grace, virtual twins used in the Boston museum [Swartout et al., 2010]. The (typical) children visiting the museum could transmit questions to a museum employee, who in turn asked the questions orally to the two virtual characters. The two virtual agents would then answer the question with animations of non-verbal behavior during their collective responses. However, this device was not dedicated to learning social skills for participants with ASD. Automatic processing and interactivity remained limited (the children themselves did not ask the questions and the two virtual agents did not change their behaviors once the animation was launched).

Numerous systems allow locating and support social interactions between two users connected via a virtual environment. For example, the Block Challenge system [Parsons, 2015] seeks to train social collaboration and being able to take the other user's point of view. It has been tested with six children with ASD and eight typical children. During this study, two children with ASD (or two typical children) were connected together and saw their respective avatars but could not see each other directly. They had to work together to manipulate colored blocks. A virtual agent represented a teacher. There is no automatic simulation or display of social interaction between the virtual characters in this system. Rather, this system focuses on allowing one of the two children to be able to take the point of view of the other child: The interactions between the two children take place in real time but are mediated by the system.

Systems for managing interactions between two co-present users and one or more virtual agents are limited. Indeed, this requires complex automatic management of the system's speech turns (knowing which user is talking about what and to whom (the other user or one of the virtual agents)). However, this type of system could be very useful in a generalization perspective to involve at the same time several people with ASD in social interactions progressively integrating other people and progressively decreasing the presence of virtual agents. The active research in this field of virtual agents allowing varied social interactions with users is however moving in this direction and is encouraging.

25.4 Similarities and Differences in IVAs and SRs

These new technologies for social training have some advantages over people: they are available at all times, they can be controlled, are never tired, ... However, new technologies also have disadvantages and potential risks (addiction, bugs, possible design flaws) that need to be assessed with experimental studies, a user-centered design approach and evaluations. This research must therefore follow the recommendations proposed for interventions concerning the learning of social skills for people with ASD [Rao et al., 2008]. Advantages may also differ for robots and virtual agents. For instance, the tangible 3D features of robots are more appealing than virtual characters for low-functioning children with ASD. The notion of presence is direct when you can touch the robot, take it in your arms or walk besides him. On the other hand, virtual animations can be more varied and simulate familiar environments which may elicit less social stress. In any case, what Scassellati and colleagues wrote in 2012 remains true today: research on robotic intervention in autism is under construction [Scasselatti et al., 2012]. This comment is true also for virtual agents.

Numerous experimental studies have investigated the recognition of emotions by participants with ASDs during non-interactive presentation of videos without situational context. For example, one study compared how participants with ASDs recognized emotions expressed in videos of people, virtual agents, and robots expressing emotions via the face and body [Chevalier et al., 2016]. The authors explored if the individual's reliance on proprioceptive and kinematic visual cues affected the way the individual suffering from ASD interacts with a social agent (human/robot/virtual agent). They assessed the potential link between recognition performances of body / facial expressions of emotion of increasing complexity, emotion recognition on platforms with different visual features (two mini-humanoid robots, a virtual agent, and a human), and proprioceptive and visual cues integration of an individual. For neurotypical individuals, the results indicate a relationship between profiles and emotion recognition. Neurotypical individuals that rely more heavily on visual cues yielded better recognition scores. However, neurotypical individuals relying on proprioception had better recognition scores. Finally, participants with ASD relying more heavily on proprioceptive cues had lower emotion recognition scores on all conditions than participants relying on visual cues.

Involving low-functioning adolescents with ASD, a study used a robotic architecture displaying facial expressions of emotions. The distribution of attention toward the mechanical and the emotional elements was analyzed through the use of eye-tracking combined with a morphing technique. It was shown that individuals with ASD process motion rather than emotional signals when facing facial expressions [Han et al., 2015].

Warren et al. (2015) programmed a robot to investigate joint attention in children with ASD. Kajopoulos et al. (2015) trained joint attention skills in 7 children with ASD using a pet robot that produced head movements to draw attention towards a target stimulus.

25.5 Current Challenges

Most of the studies using social robots and virtual agents to train people with ASD have been conducted with a limited number of participants. They also used different methodologies, some yielding quantitative measures of efficacy [Massaro & Bosseler, 2006], others relying on mostly qualitative evaluations [Tartaro, 2007].

For innovation in the field to transfer to real life, more thorough research needs to be conducted. This is crucial to build trust in the technology, especially among clinicians, family members and individuals with autism. The field therefore needs to embrace what is referred to as Evidence-

Based Practice [Mesibov & Shea, 2011], which integrates clinicians' individual expertise with the available evidence yielded by scientific research.

Applying Evidence-Based Practice to digital technology nevertheless calls for some adjustments and a specific framework has been developed for this purpose in ASD [Zervogianni et al., 2020].

25.6 Future Directions

There are still technological limitations before being able to offer social interactions in a small group that realistically, expressively and interactively involve several users and several virtual or robotic agents. Indeed, this type of situation requires a very complex automatic management of speech turns and the generation of non-verbal behaviors for the different agents in real time during the interaction with the different users. Difficult as it is however, such a multiparticipant context is highly desirable to train people with ASD to overcome their shyness and avoidance of social groups. One important element is to develop their pragmatic capacities to consider social signals coming simultaneously from different persons.

These complex challenges require a multidisciplinary approach (Computer Science, Cognitive Sciences, developmental Psychology) integrating recent advances from Artificial Intelligence and Human-Computer Interaction research for enabling realistic and credible interactions in terms of their dynamics. It is noteworthy that a multidisciplinary approach has multidisciplinary benefits. For instance, work with robots has enabled researchers to systematically explore the role of the body in shaping the development of skill, shading new light on development as a complex dynamical system [Oudeyer, 2017] and opening avenues for an enactive approach of development in ASD.

A promising avenue is to combine social skills training with other types of training, such as language or motor training. Thus, in a meta-review [Srinivasan et al., 2014], the authors draw on existing studies using exergames, for example, to propose recommendations for the integration of sports activities for obese people with autism. The studies reviewed by these authors suggest that regular physical exercise can have beneficial effects on social, behavioral, cognitive and motor dysfunctions. The authors make recommendations regarding the structure of the environment in which physical activities take place, the exercises themselves and the communication associated with the activities (instructions, feedback, reinforcement).

Another avenue that seems also relevant is to combine virtual agents with other physical interaction devices. Tangible interactions [Farr, 2011; Farr et al., 2010] are promising for people with ASD since they involve real objects and physical contact. Augmented interaction devices, which combine virtual elements (such as a virtual agent) with a physical environment, would test some generalization capabilities to physical elements present in real interactions, one of the expected characteristics of virtual technologies for social learning [Neely et al., 2016]. Mixed reality where virtual elements are added to the real physical environment is currently explored with people with ASD, as it may stimulate symbolic play capacities. Ethical precautions are however needed with people for whom the distinction between reality and symbols is not clear.

25.7 Summary

In this chapter, we have reviewed different systems and studies on the use of social robots and virtual agents to study or train the social skills of people with ASD. These technologies are promising and have already shown effective results. Robotic systems are mainly face -to-face systems, though they can be used with several users facing one robot, as in affective computing designs. For virtual agents, the different systems used involve different configurations of interactions with users. For robotic as well as for virtual systems, important improvements have been noted throughout the last twenty years. Notably, these systems tend now to be inspired by practice-based methodology instead of relying only on feasibility criteria. More and more numerous systems are the result of an interaction with developmental psychopathology which leads to get a good knowledge about neurodevelopmental specificities and include the use of follow-ups. Similarly, both robotic and virtual systems tend to use the most recent developments of Artificial Intelligence and Human-Computer Interaction to anticipate the needs of the users. Thus, the true improvements lie in a multidisciplinary approach that will be more and more used in the future.

Bibliography

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

André, E. (2002). From Simulated Dialogues to Interactive Performances with Virtual Actors. *Proceedings of the 25th Annual German Conference on AI (KI 2002),* Aachen, Germany, September 16-20.

Baylor, A., Ebbers, S. (2003). The pedagogical agent split-persona effect: When two agents are better than one. *Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications*. Honolulu, Hawai, 459-462.

Ben Youssef, A., Chollet, M., Jones, H., Sabouret, N., Pelachaud, C., & Ochs, M. (2015). An Architecture for a Socially Adaptive Virtual Recruiter in Job Interview Simulations. *Proceedings of the Second International Workshop on Intelligent Digital Games for Empowerment and Inclusion at IUI 2015*.

Billard, A., Robins, B., Nadel, J., Dautenhahn, K. (2007). Building Robota, a mini-humanoid robot for the rehabilitation of children with autism. *Assist Technol.*, 19(1):37-49.

Bono, V., Narzisi, A., Jouen, A-L, Tilmont, E., Hommel, S., …Muratori, F. and Michelangelo Study group (2016). GOLIAH: A Gaming Platform for Home-Based Intervention in Autism – Principles and Design. Frontiers in Psychology. Doi: journal.frontiers in.org/article/&0.3389/fpsyt.2016.00070.

Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., Parker, R. (2013). Academic benefits of peer tutoring: A meta-analytic review of single-case research. *School Psychology Review*, 42(1), 39-55.

Breazeal, C. Gray, J., Berlin, M. (2009). An Embodied Cognition Approach to Mindreading Skills for Socially Intelligent Robots. *International Journal of Robotics Research*, vol. 28, no. 5, pp.656-680.

Caruana, N., Stieglitz Ham, H., Brock, J., Woolgar, A., Kloth, N., Palermo, R., & McArthur, G. (2017). Joint attention difficulties in autistic adults : An interactive eye-tracking study. Autism, 1362361316676204.

Chevalier, P., Martin, J.-C., Isableu, B., Bazile, C., Tapus, A. (2016). Impact of sensory preferences of individuals with autism on the recognition of emotions expressed by two robots, an avatar, and a human. *Autonomous Robots*.

Chollet, M., Wörtwein, T., Morency, L.-P., Shapiro, A. & Scherer, S. (2015). Exploring feedback strategies to improve public speaking: An interactive virtual audience framework. *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 1143-1154.

Courgeon, M. (2011) MARC: Computer Models of Emotions and their Facial Expressions for Real-Time Affective Human-Computer Interaction. Thesis of the University of Paris Sud.

Courgeon, M., Clavel, C., Martin, J.-C. (2014). Modeling Facial Signs of Appraisal during Interaction; Impact on Users' Perception and Behavior. *Proceedings of the 13th International*

Conference on Autonomous Agents and Multiagent Systems (AAMAS'2014), Paris, France, 765–772.

Courgeon, M., Rautureau, G., Martin, J.-C., & Grynszpan, O. (2014). Joint Attention Simulation Using Eye-Tracking and Virtual Humans. IEEE Transactions on Affective Computing, 5(3), 238-250. https://doi.org/10.1109/TAFFC.2014.2335740

Courgeon, M., Rautureau, G., Martin, J.C., Grynszpan, O. (2014). Joint Attention Simulation using Eye-Tracking and Virtual Humans. *IEEE Transactions on Affective Computing*, 5 (3), 238 – 250.

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism : Social orienting, joint attention, and attention to distress. Developmental Psychology, 40(2), 271-283. https://doi.org/10.1037/0012-1649.40.2.271

DiGennaro Reed, F.D., Hyman, S.R., Hirst, J.M. (2011). Applications of technology to teach social skills to children with autism. *Research in Autism Spectrum Disorders*, 5(3), 1003-1010.

Dumas, G., Nadel, J;, Soussignan, R., Martinerie ,J. & Garnero,L. (2010). Inter-brain synchronizations during social interaction. *PlosOne*. *,* 5: doi:10.1371/journal.pone.0012166.

Ekman, P., Friesen, W.V., Hager, J.C. (2002). Facial Action Coding System: The Manual. Facial Action Coding System, user & investigator guides. A Human Face.

Emery, N. J. (2000). The eyes have it : The neuroethology, function and evolution of social gaze. Neuroscience and Biobehavioral Reviews, 24(6), 581-604. https://doi.org/10.1016/S0149-7634(00)00025-7

Farr, W., Yuill N. et Raffle H. (2010). Social benefits of a tangible user interface for children with Autistic Spectrum Conditions. *Autism*, 14(3), 237-52.

Farr, William John (2011). *Tangible user interfaces and social interaction in children with autism*. Doctoral thesis (DPhil), University of Sussex. http://sro.sussex.ac.uk/6962/

Georgescu, A., Kuzmanovic, B., Roth, D., Bente, G., Vogeley, K. (2014). The Use of Virtual Characters to Assess and Train Non-Verbal Communication in High-Functioning Autism. *Frontiers in Human Neuroscience*, 8: 807.

Grey Walter, W. (1950). An imitation of life. *Scientific American,* 42-45.

Grossard, C. & Grynszpan, O. (2015). Digital technology-assisted skills training in autism: A review. *Childhood,* 1/2015, 67-85.

Grossard, C., Grynspan, O., Serret, S., Jouen, A.-L., Bailly, K., Cohen, D. (2017). Serious games to teach social interactions and emotions to individuals with autism spectrum disorders (ASD). *Computers and Education*, 113, 195-211.

Grynszpan, O. (this volume). Haptic and tactile interfaces for autism: a systematic review. *Childhood*.

Grynszpan, O., Martin, J.-C., & Fossati, P. (2017). Gaze leading is associated with liking. Acta Psychologica, 173, 66-72. https://doi.org/10.1016/j.actpsy.2016.12.006

Grynszpan, O., Martin, J.-C., Fossati, P. (2017). Gaze Leading is Associated with Liking. *Acta Psychologica*, 173, 66–72.

Grynszpan, O., Nadel, J. (2015). An eye-tracking method to reveal the link between gazing patterns and pragmatic abilities in high functioning autism spectrum disorders. Frontiers in Human Neuroscience, 8, 1067. doi: 10.3389/fnhum.2014.01067

Grynszpan, O., Nadel, J., Constant, J., Le Barillier, F., Carbonell, N., Simonin, J., Martin, J. C., & Courgeon, M. (2009). A new virtual environment paradigm for high functioning autism intended to help attentional disengagement in a social context Bridging the gap between relevance theory and executive dysfunction. Virtual Rehabilitation International Conference, 2009, 51-58. IEEE Xplore. https://doi.org/10.1109/ICVR.2009.5174205

Grynszpan, O., Nadel, J., Martin, J.-C., Simonin, J., Bailleul, P., Wang, Y., Gepner, D., Le Barillier, F., & Constant, J. (2012). Self-monitoring of gaze in high functioning autism. Journal of autism and developmental disorders, 42(8), 1642-1650. https://doi.org/10.1007/s10803-011-1404-9

Grynszpan, O., Weiss, P. L., Perez-Diaz, F., & Gal, E. (2014). Innovative technology-based interventions for autism spectrum disorders: A meta-analysis. Autism, 18(4), 346–361.

Grynszpan,O.,Nadel, J., Martin, J.-C., Simonin, J., Bailleul, P.,Wang, Y., Gepner, D, Le Barillier, F, Constant, J. (2012). Self-monitoring of gaze in high functioning autism. *Journal of autism and developmental disorders*, 42(8), 1642-1650.

Hoque, M. E., Courgeon, M., Mutlu, B., Martin, J.-C., Picard, R. W. (2013). MACH: My Automated Conversation coacH. *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing (UBICOMP'2017)*, 697-706.

Jarrold, W., Mundy, P., Gwaltney, Bailenson, J., Hatt, N., McIntyre, N., Kim, K., Solomon, M., Novotny, S., Swain, L. (2013). Social attention in a virtual public speaking task in higher functioning children with autism. *Autism Research*, 6(5), 393-410

Kandalaft, M. R., Didehbani, N., Krawczyk, D. C., Allen, T. T., & Chapman, S. B. (2013). Virtual Reality Social Cognition Training for Young Adults with High-Functioning Autism. Journal of Autism and Developmental Disorders, 43(1), 34-44. https://doi.org/10.1007/s10803-012-1544-6

Kim, K., Rosenthal, M.Z., Gwaltney, M., Jarrold, W., Hatt, N., McIntyre, N., Swain, L., Solomon, M., Mundy, P. (2015). A Virtual Joy-Stick Study of Emotional Responses and Social Motivation in Children with Autism Spectrum Disorder. *J. Autism Dev. Disord.*, 45:3891–3899

Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. Nature, 387(6635), 767-768. https://doi.org/10.1038/42842

Kozima, H., Nakagawa, C., & Yasuda, Y. (2007). Children–robot interaction: A pilot study in autism therapy. In C. von Hofsten & K. Rosander (Éds.), Progress in Brain Research (Vol. 164, p. 385-400). Elsevier. https://doi.org/10.1016/S0079-6123(07)64021-7

Lahiri, U., Warren, Z., Sarkar, N. (2011). Design of a gaze-sensitive virtual social interactive system for children with autism. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 19(4), 443-452.

Little, G. E., Bonnar, L., Kelly, S. W., Lohan, K. S., & Rajendran, G. (2016). Gaze contingent joint attention with an avatar in children with and without ASD. 2016 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob), 15-20. https://doi.org/10.1109/DEVLRN.2016.7846780

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

Massaro, D. W., & Bosseler, A. (2006). Read my lips : The importance of the face in a computer-animated tutor for vocabulary learning by children with autism. Autism: The International Journal of Research and Practice, 10(5), 495-510. https://doi.org/10.1177/1362361306066599.

Masson, O., Baratgin, J. et Jamet, F. (2017). Nao, robot : a social clues transmitter : What impacts ? *Proceedings of the XXXth international Conference on Industrial, Engineering, Other applications of applied intelligent systems (IEA/AIE)*, 559-568.

Mesibov, G. B., & Shea, V. (2011). Evidence-Based Practices and Autism. Autism, 15(1), 114-133. https://doi.org/10.1177/1362361309348070

Milne, M., Leibbrandt, R., Raghavendra, P., Luersen, M., Lewis, T., Powers, D. (2013). Lesson authoring system for creating interactive activities involving virtual humans: the thinking head whiteboard. *Proceedings of the IEEE Symposium on Intelligent Agents*, 13-20.

Nadel, J. (2014). *How imitation boosts development in infancy and autism spectrum disorder*. Oxford, New York, NY: Oxford University Press.

Nadel, J. (2015). Perception-action coupling and imitation in autism spectrum disorder. *Developmental Medicine and Child Neurology, 57*, 21, 55-58.

Neely, L.C., Ganz, J.B., Davis, J.L., Boles, M.B., Hong, E.R., Ninci, J., Gilliland, W.D. (2016). Generalization and maintenance of functionnal living skills for individuals with autism spectrum disorders: a review and meta-analysis. *Journal of Autism and Developmental Disorders*, 3(1), 37-47.

Oudeyer, P-Y. (2017) What can we learn about development from baby robots? *WIREs Cogn Sci* 2017, 8:e1395. doi: 10.1002/wcs.1395

Parsons, S. (2015). Learning to work together: Designing a multi-user virtual reality game for social collaboration and perspective-taking for children with autism. *International Journal of Child-Computer Interaction*, 6, 28-38.

Picard, R., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., Machover, T. Resnick, M., Roy, D., Strohecker, C. (2004). Affective Learning — A Manifesto. *BT Technology Journal*, vol. 22, no. 4, pp. 253-269.

Pennisi, P., Tonacci, A., Tartarisco, G., Billeci, L., Ruta, L., Gangemi, S., Pioggia, G. (2016). Autism and Social Robotics: A Systematic Review. *Autism Res.* , 9(2):165-83.

Rao, P.A., Beidel, D.C., Murray, M.J. (2008). Social skills interventions for children with asperger syndrome or high-functionning autism: A review and recommendations. *Journal of Autism and Developmental Disorders*, 38, 353-361.

Recht, S., & Grynszpan, O. (2019). The sense of social agency in gaze leading. Journal on Multimodal User Interfaces, 13(1), 19-30. https://doi.org/10.1007/s12193-018-0286-y

Rist, T., André, E., Baldes, S., Gebhard, P., Klesen, M., Kipp, M., Rist, P., Schmitt, M. (2003). A Review of the Development of Embodied Presentation Agents and Their Application Fields. In H. Prendinger & M. Ishizuka (Ed.), *Life-Like Characters: Tools, Affective Functions, and Applications. Cognitive Technologies,* 377-404.

Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., Shah, N. J., Fink, G. R., & Vogeley, K. (2010). Minds made for sharing: Initiating joint attention recruits reward-related neurocircuitry. Journal of Cognitive Neuroscience, 22(12), 2702-2715. https://doi.org/10.1162/jocn.2009.21401

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes : An absence of spontaneous theory of mind in Asperger syndrome. Science (New York, N.Y.), 325(5942), 883-885. https://doi.org/10.1126/science.1176170

Serret, S., Hun, S., Iakimova, G., Lozada, J., Anastassova, M., Santos, A., Vesperini, S., & Askenazy, F. (2014). Facing the challenge of teaching emotions to individuals with low- and high-functioning autism using a new Serious game: A pilot study. Molecular Autism, 5(1), 37. https://doi.org/10.1186/2040-2392-5-37

Smith, M.J., Fleming, M.F., Wright, M.A., Roberts, A.G., Humm, L.B., Olsen, D., Bell, M.D. (2015). Virtual reality job interview training and 6-month employment outcomes for individuals with schizophrenia seeking employment. *Schizophr Res.*, 166(1-3):86-91.

Smith, M.J., Ginger, E.J., Wright, K., Wright, M.A., Taylor, J.L., Humm, L.B., Olsen, D.E., Bell, M.D. & Fleming, M.F. (2014). Virtual reality job interview training in adults with autism spectrum disorder. *Journal of Autism Developmental Disorders*, 44(10):2450-63.

Smith, M.J., Humm, L.B., Fleming, M.F., Jordan, N., Wright, M.A., Ginger, E.J., Wright, K., Olsen, D., Bell M.D. (2015). Virtual Reality Job Interview Training for Veterans with Posttraumatic Stress Disorder. *J Vocat Rehabil.*, 42(3):271-279.

Srinivasan, S. M., Pescatello, L. S., Bhat, A. N. (2014). Current Perspectives on Physical Activity and Exercise Recommendations for Obesity and Physical Fitness in Children and Adolescents With Autism Spectrum Disorders. *Physical Therapy*, 94(6).

Swartout, W., Traum, D., Arstein, R., Noren, D., Debevec, P., Bronnenkant, K., Williams, J., Leuski, A., Narayanan, S., Piepol, D., Lane, C., Morie, J., Aggarwal, P., Liewer, M., Chiang, J.Y., Gerten, J., Chu, S., White, K. (2010). Ada and Grace: Toward Realistic and Engaging Virtual Museum Guides. Proceedings of the International Conference on Intelligent Virtual Agents (IVA'2010), LNAI 6356, Springer, 286-300.

Tartaro, A. (2007). Authorable virtual peers for children with autism. *Proceedings of the extended abstracts of the International Conference on Human Factors in Computing Systems (CHI'2007)*, 1677-1680. ACM New York.

Wiener, N. (1948). *Cybernetics or control and communication in the animal and the machine*. Cambridge, Mass: the MIT Press.

World Economic Forum (2016). New Vision for Education: Fostering Social and Emotional Learning through Technology. Mars 2016 http://www3.weforum.org/docs/WEF_New_Vision_for_Education.pdf

Yun, S., Choi, J., Park, S., Bong, G., Yoo, H. (2017). Social skills training for children with autism spectrum disorder using a robotic behavioral intervention system. *Autism. Res*.

Zervogianni, V., Fletcher-Watson, S., Herrera, G., Goodwin, M., Pérez-Fuster, P., Brosnan, M., & Grynszpan, O. (2020). A framework of evidence-based practice for digital support, co-developed with and for the autism community. Autism, 1362361319898331. https://doi.org/10.1177/1362361319898331