Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Attention shifting during child–robot interaction: a preliminary clinical study for children with autism spectrum disorder*#

Guo-bin WAN[1], Fu-hao DENG[2], Zi-jian JIANG[2], Sheng-zhao LIN[2], Cheng-lian ZHAO[2],
Bo-xun LIU[2], Gong CHEN[2], Shen-hong CHEN[3], Xiao-hong CAI[4],
Hao-bo WANG[5], Li-ping LI[6], Ting YAN[7], Jia-ming ZHANG[†‡2]

[1]Shenzhen Maternal & Child Healthcare Hospital, Shenzhen 518000, China
[2]Institute of Robotics and Intelligent Manufacturing, The Chinese University of Hong Kong (Shenzhen), Shenzhen 518172, China
[3]School of Mechanical Engineering, Zhejiang University, Hangzhou 310027, China
[4]Institute of Information Engineering & Technology, Huaqiao University, Xiamen 361021, China
[5]School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China
[6]School of Mechanical and Electrical Engineering, Beijing Information Science and Technology University, Beijing 100192, China
[7]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[†]E-mail: zhangjiaming@cuhk.edu.cn

Received Sept. 11, 2018; Revision accepted Jan. 28, 2019; Crosschecked Mar. 14, 2019

Abstract: There is an increasing need to introduce socially interactive robots as a means of assistance in autism spectrum disorder (ASD) treatment and rehabilitation, to improve the effectiveness of rehabilitation training and the diversification of treatment, and to alleviate the shortage of medical personnel in mainland China and other places in the world. In this preliminary clinical study, three different socially interactive robots with different appearances and functionalities were tested in therapy-like settings in four different rehabilitation facilities/institutions in Shenzhen, China. Seventy-four participants, including 52 children with ASD, whose processes of interacting with robots were recorded by three different cameras, all received a single-session three-robot intervention. Data were collected from not only the videos recorded, but also the questionnaires filled mostly by parents of the participants. Some insights from the preliminary results were obtained. These can contribute to the research on physical robot design and evaluations on robots in therapy-like settings. First, when doing physical robot design, some preferential focus should be on aspects of appearances and functionalities. Second, attention analysis using algorithms such as estimation of the directions of gaze and head posture of a child in the video clips can be adopted to quantitatively measure the prosocial behaviors and actions (e.g., attention shifting from one particular robot to other robots) of the children. Third, observing and calculating the frequency of the time children spend on exploring/playing with the robots in the video clips can be adopted to qualitatively analyze such behaviors and actions. Limitations of the present study are also presented.

Key words: Human–robot interaction; Robot-enhanced therapy; Socially interactive robots; Robot-mediated intervention
https://doi.org/10.1631/FITEE.1800555 **CLC number:** TP242.6; R748

1 Introduction

1.1 A brief introduction to autism and autism healthcare in China

According to "Autism: a global framework for action" (Munir et al., 2016), it was estimated that, worldwide, approximately 52 million people live with autism spectrum disorder (ASD), and 86.5% of all cases of ASD had been reported in high-income

‡ Corresponding author
* Project supported by the Shenzhen Science and Technology Innovation Commission, China (Nos. JCYJ20170410172100520 and GJHZ20160229200136090)
# Electronic supplementary materials: The online version of this article (https://doi.org/10.1631/FITEE.1800555) contains supplementary materials, which are available to authorized users
ORCID: Jia-ming ZHANG, http://orcid.org/0000-0001-9146-0501
© Zhejiang University and Springer-Verlag GmbH Germany, part of Springer Nature 2019

countries (HICs), where only 20% of the world population resided in 2016. As for China, a meta-analysis of 18 published suitable studies from 1987 to 2011 estimated that the prevalence of childhood autism was 11.8 per 10 000 individuals in mainland China, and of autism spectrum conditions (ASC) was 26.6 per 10 000 in mainland China, Hong Kong, and Taiwan (Sun et al., 2013). In 2017, according to WUCAILU ASD Research Institute (2017), the prevalence of ASC was also 1% in mainland China as opposed to that in HICs, based on its 10-year long intervention data collected from more than 6000 autistic children from 47 autistic institutions in 29 provinces of mainland China. Consequently, WUCAILU estimated that there were over 10 million people living with ASD and over two million with childhood autism in mainland China, with an increased population of 0.2 million people living with ASD every year. For rehabilitation training, according to Zheng (2017), a survey showed that by the end of September 2016, there were 1345 autism rehabilitation institutions registered in China Disabled Persons' Federation, nearly 50% of which were established by parents, especially the parents of autistic children, resulting in difficulty in guaranteeing a good training standard and a good rehabilitation effect. Furthermore, the number of autistic children who can receive formal rehabilitation training is still a minority. China's actual autism rehabilitation teachers can cover only 1.3% of the ASD population, and the remaining 98.7% cannot obtain effective rehabilitation training. This results mainly from a lack of qualified rehabilitation training teachers, a long reservation time for a place in the center, and the rejection of autistic children with no citizenship by public autism rehabilitation institutions.

To sum up, there is a large population of autistic people, and lack of effective rehabilitation training and qualified professionals in this field for autistic people. These are the main challenges for ASD treatment and rehabilitation in mainland China currently. Consequently, to improve the effectiveness of rehabilitation training and the diversification of treatment, and to alleviate the shortage of medical personnel, the need to introduce socially interactive robots as a means of assistance in ASD treatment and rehabilitation has become stronger for medical research in HICs and China.

1.2 Why robot-enhanced therapy in autism?

Can social robots be a useful tool in autism therapy? A meta-analysis of 861 studies shows that robot-enhanced therapy (RET) can improve the performance on three levels (behavioral, cognitive, and subjective) taken together (Costescu et al., 2014). Robots provide therapists and researchers an easy way to connect autistic subjects: e.g., ASD subjects often perform better with a robot partner rather than a human partner; ASD subjects show reduced repetitive and stereotyped behaviors and improved spontaneous language during therapy sessions, based on a systematic literature review of the studies on RET that were published in the last 10 years (Pennisi et al., 2016). Using robots with supervised autonomy in therapy for children with ASD is encouraged; however, replacing therapists by robots should be avoided (Coeckelbergh et al., 2016).

To promote social robots as useful tools in autism therapy, four world-leading research groups, the Kerstin Dautenhahn Group (Robins and Dautenhahn, 2014; Wainer et al., 2014; Huijnen et al., 2017), the Ayanna Howard Group (English et al., 2017; Lee et al., 2017), the Maja Matarić Group (Greczek and Matarić, 2015; Matarić, 2017; Clabaugh et al., 2018), and the Bram Vanderborght Group (Simut et al., 2016; Esteban et al., 2017), have been doing pioneering work.

1.3 Is it clinically useful with robot-enhanced therapy?

Robotics for autism can achieve the following targets: assisting the diagnostic process, improving eye contact and self-initiated interactions, turn-taking activities, imitation, emotion recognition, joint attention (JA), and triadic interactions (Pennisi et al., 2016). Robots can potentially be applied to 24 of 74 ASD objectives in eight domains, including sensory experiences and coping, social/interpersonal interactions and relations, functioning in daily reality, emotional wellbeing, communication, play, motor experiences and skills, and preschool skills for children with ASD (Huijnen et al., 2016). However, making robots and human–robot interaction (HRI) useful for autism intervention in clinical settings has made minimal progress in advancing robots as clinically useful for ASD intervention (Begum et al., 2016). Diehl et al. (2012) performed a critical review of the contemporary literature on the clinical use of robots in ASD therapy and diagnosis, and concluded that most studies reviewed focus on technology development rather than clinical application, and the majority of studies are exploratory and have methodological limitations. Nevertheless, Diehl et al. (2012) enumerated four categories for clinical applications of interactive robots: (1) the response of individuals with ASD to robots or robot-like behavior in comparison to human behaviors, (2) the use of robots to elicit behaviors, (3) the use of robots to model, teach, and/or practice a skill, and (4) the use of robots to provide feedback on performance. Moreover, it is necessary to clarify whether sex, intelligence quotient, and age of participants affect the outcome of therapy and whether beneficial effects occur during only the robotic session or they are still observable outside the clinical/experimental context (Pennisi et al., 2016).

1.4 Objectives of this study

Though the field of socially interactive robots for the treatment and study of autism disorders contains many studies with different methods and goals, as roboticists who had worked closely with autism researchers for more than a decade, Scassellati et al., (2012) divided it into three connected but discrete phases, i.e., physical robot design, HRI design, and evaluations of robots in therapy-like settings. They elaborated that physical robot design involves creating a physical robot and addresses many questions about its appearance and functionality, emphasizing the degree of anthropomorphism, the size of the robot, etc. HRI design involves the design of the robot's behavior when it interacts with a person, taking into account whether a robot's behaviors can be adapted to an individual's preferences and moods, and whether a robot's behaviors can be learned from the interaction or they are prespecified. Evaluations of robots are the test of a robot's physical and interaction designs in therapy-like settings, and focus on using one-time interactions or multiyear longitudinal studies, using evaluation metrics that range from qualitative behavioral analysis to quantitative measures such as time spent on performing prosocial actions, etc. For physical robot design, there is a variation in form and a function of robots for autism, as most research groups design their own robots resulting from a few commercially available robot platforms suitable for autism therapy research (Scassellati et al., 2012). Scassellati et al. (2012) categorized them by the level of anthropomorphism, into humanoid, animal-like, and machinelike (nonbiomimetic) systems. Similarly, Pennisi et al. (2016) made a classification of robots based on aesthetic characteristics of humanoid, animal-like, and non-humanoid. However, only a few robots with a touch screen were included in both of these two reviews. Future studies should investigate whether robots with a touch screen, such as the SoftBank robot Pepper, are suitable for RET for autism.

For evaluations of robots in therapy-like settings, regarding the way in which data are analyzed, most socially assistive robotic (SAR) studies involved qualitative reports of robot effects, and some studies extracted quantitative data from behavioral observation (Scassellati et al., 2012). However, these studies usually performed quantitative data analysis by recording videos of the interactions and analyzing these videos using a coding scheme of second-by-second analysis. Scassellati et al. (2012) argued that such video coding involves significant time and effort particularly because it requires training on the coding schema and a validation that the coding performed reliably. Therefore, future studies should apply more reliable and robust methods for automated video coding in autism data analysis. For robot-mediated intervention (RMI), Begum et al. (2016) argued that most behavioral data from HRI studies are collected through the subjective observation of a human being, whether done in real time or through the observation of video-recorded data (also known as "behavioral coding"). Thus, the need for obtaining a second and independent set of observed data becomes important.

In this study, we focus on two of the three aspects of designing socially interactive robots for ASD therapy, i.e., physical robot design and evaluations of robots in therapy-like settings, by investigating the following questions: (1) For physical robot design, what are the preferences of children with ASD and their parents for appearances and functionalities of the robots during the interaction? (2) For evaluations of robots in therapy-like settings, how to qualitatively analyze and quantitatively measure the prosocial behaviors and actions performed by children with ASD during the interaction?

Contributions of this study are two-fold. First, for physical robot design, we investigate whether robots with a touch screen are suitable for RET for autism by using three different robots which all have a touch screen but differ in appearance and functionality. Results obtained from simple statistical analysis of questionnaires and video analysis can provide physical robot design guidelines for appearance and functionality for commercial robot platforms. Second, for robots in therapy-like settings, we apply a more reliable, robust, and objective method for quantitative data analysis by adopting algorithms such as face detection and filtering and estimation of the directions of gaze and head posture for automated attention analysis based on recorded videos.

2 Study design

2.1 Procedure

Children with ASD or developmental delay (DD) were invited to receive a single-session three-robot intervention in four different rehabilitation facilities/institutions mostly for children with ASD in Shenzhen, China, from January 22 to 31, 2018 except the weekends.

All intervention sessions were delivered by at least one operator who is familiar with the three-robot operation and one facilitator who can instruct the child or parent(s) to go through the session. An additional healthcare worker provided a facilitation and assistance to our facilitators in each institution. Three different socially interactive robots were tested in therapy-like settings. They differ in appearance and functionality (Fig. S1 and Table S1 in the supplementary materials). The procedure of HRI in each session was designed as illustrated in Fig. 1.

In Fig. 1, there are 12 steps in each session including three steps for greeting. For greeting, each robot will raise one hand and say "大家好，我是 XXX（机器人的名字），很高兴和你做朋友" (translated into English as "Hi everyone, I am XXX (name of the robot), glad to be your friend"). Each session can be divided into two parts: in the first part, each child was told to only stay sitting in chair to watch the three robots performing (e.g., greeting, singing without body movements, telling a story, and singing with dancing movements), while in the second part each child was encouraged to explore/play with the functionalities of the robots. Each session took on average 10 min. The first part lasting approximately 5.5 min, and the duration of the second part depends on how willing each child was to play with the three robots. It is necessary to mention that when the three robots were playing music or a story video at the same time, only one of them could play with sound and the others

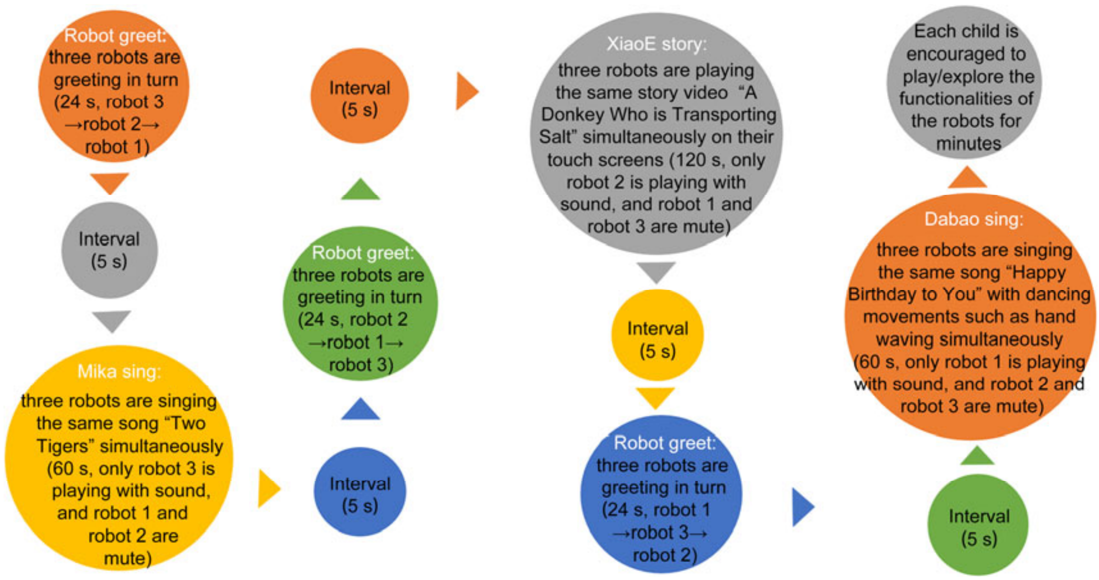

Fig. 1 Procedure of human–robot interaction (Dabao: robot 1; XiaoE: robot 2; Mika: robot 3)

were mute. We did this just to avoid confusion and annoyance for children in the interaction, as some of them might be very sensitive to multiple sound sources.

2.2 Study setups and participants

The four different rehabilitation facilities/ institutions are ASD Rehabilitation Training Center for Children, Shenzhen Maternal & Child Healthcare Hospital (MCH), and three branches of Zi Fei Yu Rehabilitation Training Center for Autistic Children (i.e., Futian, Meilin, and Longhua branches), Shenzhen.

To qualitatively analyze and quantitatively measure the prosocial behaviors and actions performed by children during the interaction, three cameras were placed in three different positions in a quiet room in each institution, recording the process of children interacting with robots from different angles (Fig. 2).

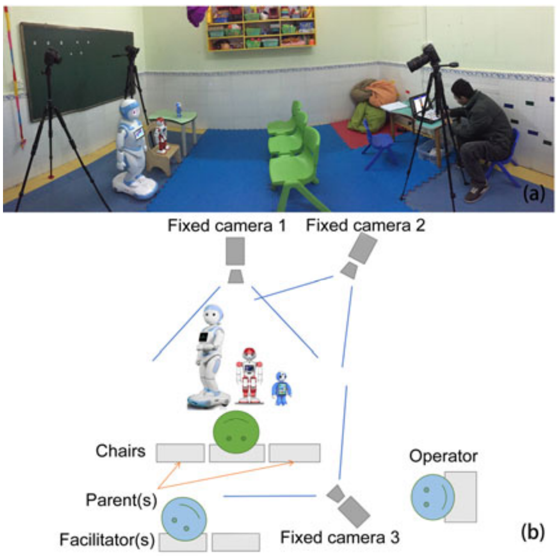

Fig. 2 Study setup in one of the four different rehabilitation facilities (a) and an example of how cameras were set in a quiet room (b)
The child sitting in the middle chair is indicated with green color with parent(s) sitting next to him/her. References to color refer to the online version of this figure

There were a total of 74 participants (63 boys and 11 girls in an average age of five years and eight months, who received average rehabilitation time of 29 months before this study) enrolling in this preliminary clinical study. Fifty-two of them were ASD children and 18 were DD children, and the rest were three not-yet-diagnosed (NYD) children and one typical development (TD) child.

2.3 Data collection

We collected data from two types of materials. One is the questionnaires filled in mostly by parents except a few professionals who accompanied the participants in the whole session, and the other is the videos recorded by the three cameras. The video recorded during each session can be divided into two parts, namely the "watching part" in which each child only watched the robots performing, and the "exploring part" in which each child was encouraged to explore the functionalities of the robots (Fig. 1). In each session, parents/professionals were asked to fill in a questionnaire in Chinese (see the supplementary materials) along with a consent form for video recording. Some of them filled in the questionnaires during the watching part, while others during the exploring part. The questionnaires and the videos collected can serve many purposes. Some examples are elaborated in Table 1.

2.4 Data analysis

2.4.1 Analysis of questionnaire data

As the questionnaires were mostly filled by parents (we treated a few professionals who accompanied the participants in the sessions as parents for simpler statistical analysis) with either ASD or DD children, we first statistically analyzed 44 questionnaires filled in by parents with ASD children, and then analyzed 18 questionnaires filled in by parents with DD children. Then we visualized and compared these statistical analyses in diagrammatic representation.

2.4.2 Video analysis of the watching part

We analyzed the watching part of the videos (recorded mostly from fixed camera 1) using attention analysis. Fig. 3 is a flowchart of attention analysis. The whole attention analysis can be divided into three parts: face detection and filtering, OpenFace attention analysis, and visualization of the attention data. Each part will be explained in detail as follows:

1. Face detection and filtering

For each sample video, it is necessary to detect the face of the participant. In the video, faces of the

Table 1 Effective samples and purposes of questionnaires and videos

Materials		Effective sample	Purpose
Questionnaires		Totally 63: 44 ASD children, 18 DD children, and one TD child	To seek insights on how to design socially interactive robots for ASD therapy that can fit the preferences of parents with ASD child/children for appearances and functionalities of the robots
Videos	Watching part	Totally 56: 40 ASD children, 14 DD children, and two TD children (one sibling accompanying her little brother was not counted as a participant, whose video was effective in this part)	To perform attention analysis using software to automatically detect each child's gaze and head pose direction in video clips, to find out which robot and what functionalities of the robot interested the children with ASD
	Exploring part	Totally 70: 50 ASD children, 18 DD children, one TD child, and one NYD child	By observing the video clips to calculate how interested, for how much time, and what functionalities the children played with the three robots

ASD: autism spectrum disorder; DD: developmental delay; TD: typical development; NYD: not-yet-diagnosed

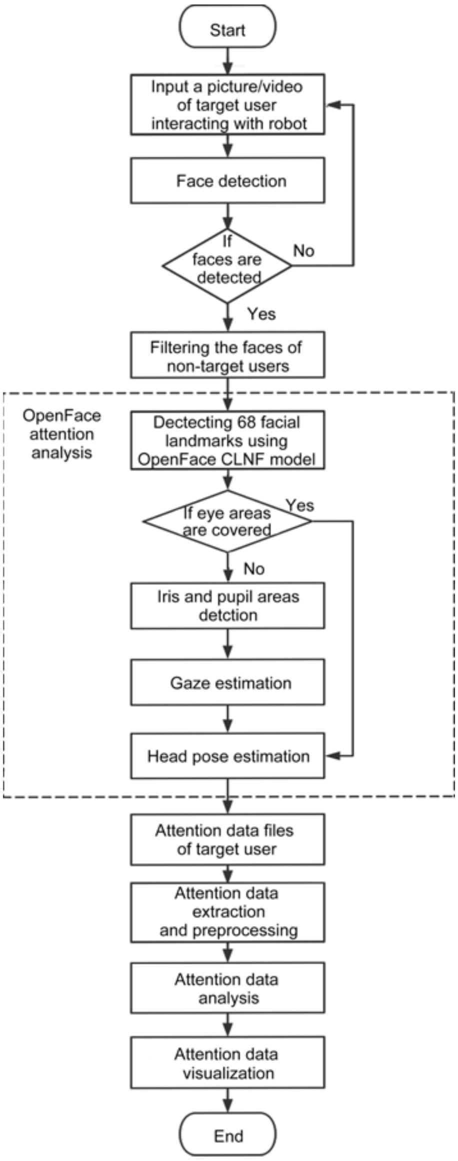

Fig. 3 Flowchart of attention analysis
CLNF: conditional local neural field

parents or experimenters in the background are often seen. These non-target user faces would greatly interfere with the analysis results. Therefore, after performing face detection, it is necessary to filter the faces of non-target users, as illustrated in Fig. 4a.

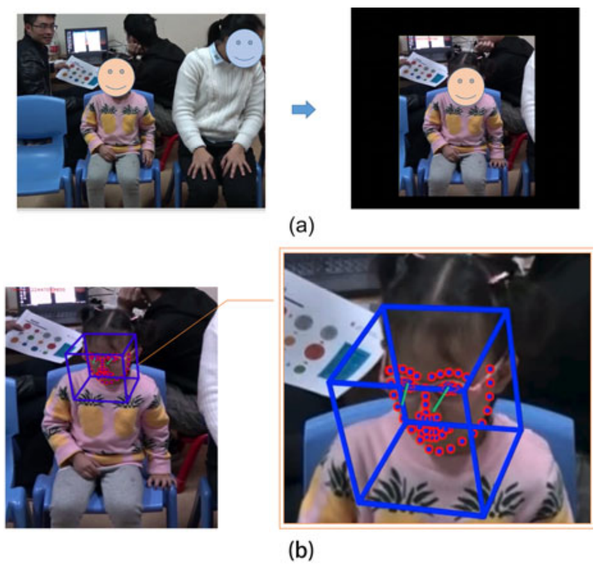

Fig. 4 Non-target user filtering (a) and attention analysis of the targeted user (b)
Red dots represent the facial feature points, blue frame indicates the estimated head posture direction, and green line indicates the estimated gaze direction. References to color refer to the online version of this figure

2. OpenFace attention analysis

Using the face area of the child participant obtained in the previous step, the conditional local neural field (CLNF) model for facial landmark detection and tracking in OpenFace (Baltrušaitis et al., 2016) can detect 68 facial feature points in the face area. These facial landmarks include facial contours, eyebrows, nose, lips, and eyes. The positions of the 68 facial landmarks and the position and parameters of the camera can be used to estimate the head posture of three rotation angles (raw, pitch, and yaw) relative to the camera position of the head of the participant.

In particular, if the eye area of the face is not obscured, then OpenFace can further detect 16 iris and 16 pupil landmarks in the eye area to do gaze estimation of two rotation angles (pitch and yaw) relative to the camera position of the eyes of the child participant. An example of using OpenFace to estimate the directions of the head and gaze of a child is illustrated in Fig. 4b. The attention of the child participant can be judged by the angle of head attitude and eye sight angle estimated by OpenFace.

3. Visualization of the attention data

First, for each video clip, OpenFace will output a corresponding attention data file, which contains each frame number, timestamp, face detection confidence, face detection success flag, two-dimensional (2D) positions of 68 facial landmarks, 2D positions of 56 eyes area landmarks, two gaze estimation angles, and three head posture estimation angles. Then, for each file, the average line angles of the head posture and gaze (the radian angles from the original file will be transformed into line angles) and the attention angle in every second are obtained.

Second, the range of the attention angle of the child participant corresponding to the particular robot to which the child participant was actually attracted can be obtained. In other words, when the attention angle of a child participant fits in a certain range, it can indicate which particular robot the child participant actually focused on (Table 2).

Third, the attention data file can be used to calculate the time that each child participant spent on paying attention to different robots. First, for each time slot, by applying the mapping in Table 2, the name of the particular robot attracting the child can be obtained. Second, as there are six different kinds of sequential robotic behaviors (Fig. 1), the time each child participant spent on paying attention to each of the three robots can be calculated. As an example, for the attention data file of the child participant labeled as sample No. 1 of MCH hospital (MCH-No.1) (Fig. 5), in the first 24 s, the child participant was focusing on robot 1 (Dabao) for 1 s, robot 2 (XiaoE) for 14 s, and robot 3 (Mika) for 9 s. In the next 60 s, the child participant was attracted to Dabao for 9 s, XiaoE for 36 s, Mika for 9 s, and for 6 s the child participant did not pay attention to any of the three robots.

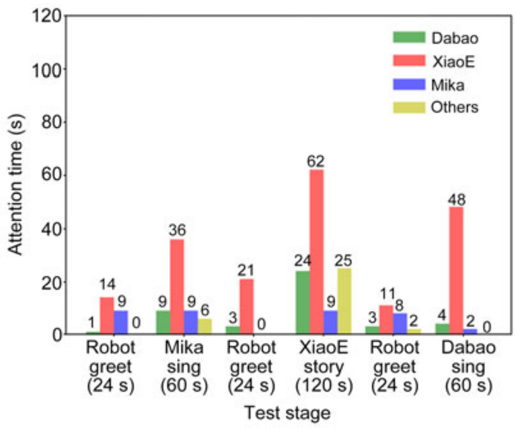

Fig. 5 Attention analysis of sample No. 1 of Shenzhen Maternal & Child Healthcare Hospital

Therefore, by observing the difference between the time each child participant spent on paying attention to different robots, it can be inferred whether a child participant focused on a particular robot more than other robots or whether one preferred some certain functionalities when the sequential robotic behaviors were performed during the whole watching part of the video, to some extent.

Table 2 Relationships between the range of the attention angle of a child participant and the particular robot attracting the child in four different institutions

Trial institution		Robot attracted to the child participant with different attention angles			
Institution	Branch	10°–40°	−10°–10°	−40°–−10°	Other
ASD center of Shenzhen Maternal & Child Healthcare Hospital	–	Dabao	XiaoE	Mika	The child participant is not attracted by any robot
Zi Fei Yu Rehabilitation Training Center for Autistic Children	Futian	Mika	XiaoE	Dabao	
	Meilin	Mika	XiaoE	Dabao	
	Longhua	Mika	XiaoE	Dabao	

ASD: autism spectrum disorder; Dabao: robot 1; XiaoE: robot 2; Mika: robot 3

2.4.3 Video analysis of the exploring part

In this part, we do video analysis by observing how each participant interacted with the robots in the video clips (recorded mostly from fixed camera 2 and camera 3). We would like to know, after watching the robots performing some actions, whether or not children were interested in interacting with the robots; if they did, which of the three robots were they interested in most? How did they interact and what functionalities of the robots did they play with most? To calculate the degree of preference of children with ASD or DD for playing with each of the three robots, we first define some parameters in Table 3.

In some sessions, parents were observed guiding their children to interact with the robots. For example, some parents might find robot 1 (the tallest one) more interesting, and they would suggest their children play with it. We define the behavior that a child interacted with a particular robot being held in his/her parent's arm and guided by the parent as strongly guided behavior (A), and the behavior that a child listened to his/her parent's guiding while sitting on a chair as weakly guided behavior (B). To balance the influence of parent's guidance on children's preferences for the robots, we multiply the relative and absolute preferences for robot 1, robot 2, and robot 3 by W_A=0.5 or W_B=0.75, respectively, depending on whether children's preferences were influenced by strongly guiding (A) or weekly guiding (B).

Table 3 Parameters on the degree of preference of children with ASD or DD for playing with robots

Parameter	Description
T	Time given to each child to explore/play with the functionalities of the three robots (the facilitator(s) instructed the parents to start or to end exploring before the parents instructed their children)
T_e	Time that each child actually spent on exploring/playing with the three robots, i.e., time that each child spent on touching, operating, and imitating one particular robot, or observing the robot at a very close distance, etc. (Fig. 6) during T
T_1	Time each child actually spent on exploring/playing with robot 1 (Dabao) during T_e
T_2	Time each child actually spent on exploring/playing with robot 2 (XiaoE) during T_e
T_3	Time each child actually spent on exploring/playing with robot 3 (Mika) during T_e
P	The score/degree of preference of children with ASD or DD for playing with the three robots: $$P=100T_e / T.$$
P_{1r}	The relative score/degree of preference of children with ASD or DD for playing with robot 1: $$P_{1r}=100T_1 /T_e.$$
P_{1a}	The absolute score/degree of preference of children with ASD or DD for playing with robot 1: $$P_{1a}=P_{1r}P /100.$$

ASD: autism spectrum disorder; DD: developmental delay. Similarly, the relative and absolute preferences for robot 2 and robot 3 can be expressed as (P_{2r}, P_{2a}) and (P_{3r}, P_{3a}), respectively

3 Preliminary results

3.1 Insights of questionnaire analysis

3.1.1 For questionnaires filled by parents with ASD children

For simple statistical analysis of 44 questionnaires filled in by parents with ASD children, we have the following insights from four aspects:

1. About parents and children with ASD: 79.55% of the parents taking care of autistic children are female; 86.36% of children in autistic children are male; and 75.00% of autistic children are 3–8 years old.

2. About the appearances of children companion robots: Parents are more concerned about the safety and durability of robots. They are satisfied with the

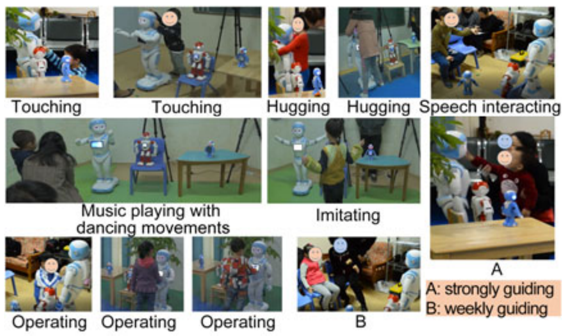

Fig. 6 Examples of how children with ASD/DD interacted with the robots in the exploring part of videos
ASD: autism spectrum disorder; DD: developmental delay

three robots' materials, walking methods, sizes, and colors. They are satisfied most with robot 1 (Dabao). On average, 63.64% of the parents hope that the color of robot is bright and warm, 50.00% of the parents expect the robot material to feel smooth, and 47.73% of the parents expect the robot's gender image to be male.

3. About the functionalities of children companion robots: On average, 63.63% of the parents expect the robots to be enthusiastic, smart, and cute, and do not want the robot to be naughty and cool, 88.64% of the parents hope that robots can accompany their children to play, and 71.21% of the parents expect the robot to give guidance in training under certain conditions and have a certain therapeutic effect on ASD.

4. About the acceptance of children companion robots: 84.09% of the parents accept the robots and are happy to recommend these robots to others, and 81.81% of the parents want the price of the robots to be lower than RMB 5000.

3.1.2 For questionnaires filled by parents with DD children

For simple statistical analysis of 18 questionnaires filled in by parents with DD children, we have the following insights from four aspects:

1. About parents and children with DD: The percentage of female parents taking care of children with DD is 83.33%, 88.89% of children with DD are male, and 77.78% of DD children are 3–8 years old.

2. About the appearance of children companion robots: Parents are more concerned about the safety and durability of robots. They are satisfied with the three robots' materials, walking methods, sizes, and colors. They are satisfied most with robot 1 (Dabao). On average, 50.00% of the parents hope that the color of the robot is bright and warm, 66.67% of the parents expect the robot material to feel smooth, and 38.89% of the parents expect the robot's gender image to be male.

3. About the functionalities of children companion robots: On average, 61.11% of the parents expect the robots to be enthusiastic, smart, and cute, and do not want the robot to be naughty and cool; 88.89% of the parents hope that robots can accompany their children to play; 64.81% of the parents expect the robot to give guidance training under certain conditions and have a certain therapeutic effect on DD.

4. About the acceptance of children companion robots: 77.78% of the parents accept the robots and are happy to recommend these robots to others; 72.22% of the parents want the price of robots to be lower than RMB 5000.

3.1.3 Comparison between the two types of questionnaires

In comparison, as can be seen from the simple statistical analysis of the two types of questionnaires, most of children with ASD or DD are male, and most of the parents tending them are female. Furthermore, parents with ASD or DD children have more or less the same expectation on the appearances, functionalities, and acceptance of children companion robots.

3.2 Insights of video analysis: the watching part

There were a total of 74 participants whose processes of interacting with robots were recorded by three cameras in the therapy-like settings. However, when performing video analysis of the watching part, it was found that due to various reasons in each session, such as that some child participants did not want to cooperate with the facilitator, or some participants behaved paralyzed sitting, walking around, or sloshing their bodies too often, it resulted in a greatly reduced recording quality and analyzability of the videos, and more difficulties in using the OpenFace algorithm to analyze these samples continuously and stably. Thus, the final effective samples for attention analysis and visualization were reduced to 40 ASD children, 14 DD children, and two TD children.

With these 56 effective samples of videos, attention analysis for each sample or for the three different kinds of samples (i.e., ASD, DD, and TD children) can be visualized using bar charts. Bar graphs with variance for average attention of each kind of sample are illustrated in Fig. 7. Then detailed statistical analysis can be done using SPSS (Statistical Packages for the Social Sciences) version 19.

When watching the robots performing the six sequential robotic behaviors, ASD children shared similar patterns with DD children (Fig. 8). After performing an SPSS variance analysis (general linear model to multivariate) of the attention time of 40 ASD children samples and 14 DD children samples, no significant difference was found in the attention of the two different groups to the three different kinds of robots.

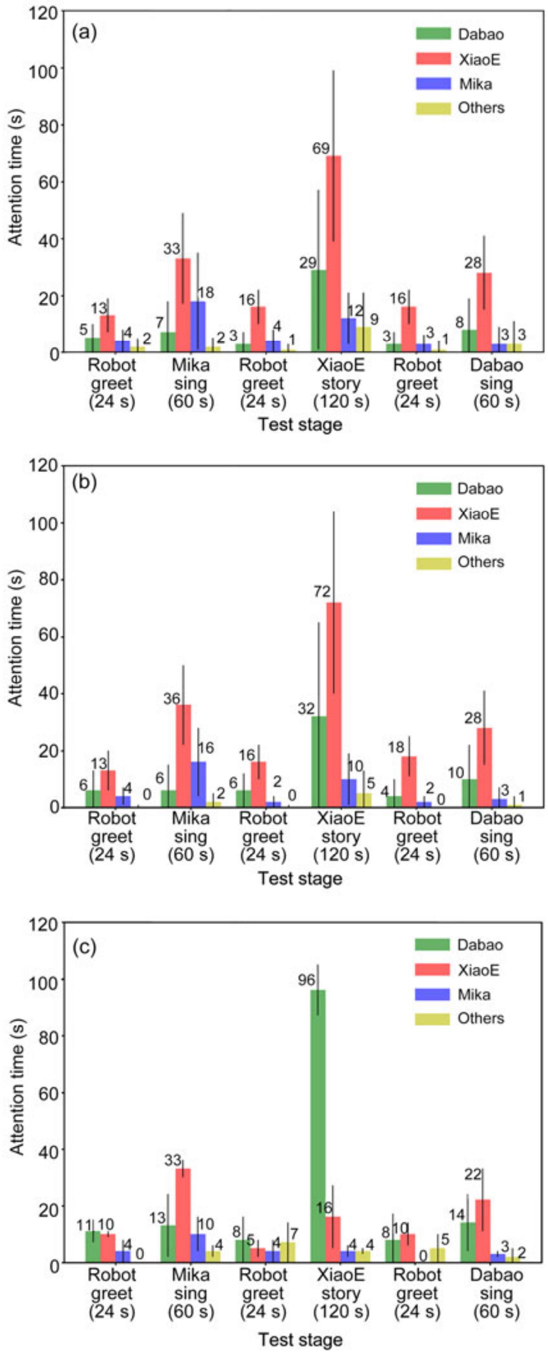

Fig. 7 Average attention bar graph with variance for 40 ASD children samples (a), 14 DD children samples (b), and two TD children samples (c)
ASD: autism spectrum disorder; DD: developmental delay; TD: typical development. References to color refer to the online version of this figure

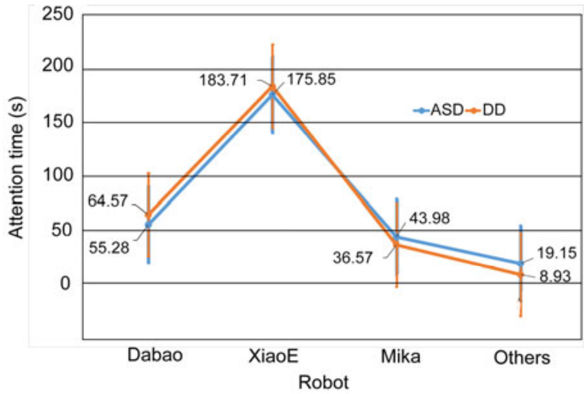

Fig. 8 Patterns in paying attention to different robots for ASD children and DD children
ASD: autism spectrum disorder; DD: developmental delay. References to color refer to the online version of this figure

For ASD children, they focused on robot 2 (XiaoE) more than on robot 1 (Dabao) and robot 3 (Mika) for approximately 295 s (by adding all the duration time of the six sequential robotic behaviors together). In SPSS variance analysis, the general linear model was used to repeated measures, where the mean differences are significant as $p_1=p_2=0$. Furthermore, robot 1 (Dabao) drew a little more attention from the children than robot 3 (Mika) did; however, there was no significant difference (p=0.291).

For DD children, they focused on robot 2 (XiaoE) more than on robot 1 (Dabao) and robot 3 (Mika) for approximately 295 s. In SPSS variance analysis, the general linear model was used to repeated measures, where the mean differences are significant as p_1=0.006 and p_2=0.000. Furthermore, robot 1 (Dabao) drew a little more attention from the children than robot 3 (Mika) did; however, there was no significant difference (p=0.207).

In other words, in the watching part, children with ASD or DD had focused on robot 2 (XiaoE) for most of the time. This might due to the fact that robot 2 (XiaoE) was placed in the middle of the settings, with robot 1 (Dabao) and robot 3 (Mika) being on either side of robot 2 (XiaoE) (Fig. 2), resulting in the fact that robot 2 (XiaoE) could draw attention from children more easily than robot 1 (Dabao) and robot 3 (Mika) did for each of the six sequential robotic behaviors. Consequently, the children naturally spent more time in watching the robot in the middle. This is consistent with the recent finding that people with ASD have a stronger image center bias regardless of object distribution, even there is no object in the center (Wang et al., 2015).

As there were only two subjects involved in the TD condition, the number was not sufficient for drawing solid conclusions. Consequently, results from only two TD subjects were not included in comparisons with statistical results of ASD or DD children.

As for children's preferences for some certain functionalities of the three robots, for children with ASD or DD, when the robots were performing the six sequential robotic behaviors, they spent most of the time watching the robots, indicating that they were interested in all of these functionalities (the percentage of the overall time nodes of "others" in each of the six sequential robotic behaviors was low).

3.3 Insights of video analysis: the exploring part

Degrees of preference of children with ASD or DD for playing with each of the three robots are calculated in Tables 4 (without considering the influence of parental guidance) and 5 (considering the influence of parental guidance).

As indicated by Table 4, without considering the influence of parents, for children with ASD or DD, robot 1 is more attractive than robots 2 and 3; i.e., children spent more time playing with robot 1, and robot 2 was more or less the same attractive as robot 3 for children with ASD. However, for children with DD, robot 3 is more attractive than robot 2.

As indicated in Table 5, whether or not considering the influence of parental guidance, for children with ASD or DD, robot 1 is more attractive than robots 2 and 3.

Consequently, among all these three robots with different sizes, colors, appearances, etc., robot 1 is the most attractive. To investigate what functionalities of the robots are more attractive, we also observed how children with ASD/DD interacted with the robots (in Fig. 6 as examples) that they played with most, and calculated the frequency (Table 6).

Table 4 Scores of relative and absolute preferences for robots 1, 2, and 3 without considering the influence of parental guidance

Sample	Relative (P_{nr}) and absolute (P_{na}) preferences for the three robots (n=1, 2, 3)					
	Robot 1		Robot 2		Robot 3	
	P_{1r}	P_{1a}	P_{2r}	P_{2a}	P_{3r}	P_{3a}
50 ASD children	60.3	42.1	15.0	10.6	16.6	9.8
18 DD children	54.3	32.8	16.5	10.1	23.8	15.6
One TD child	23.4	21.2	22.4	20.3	54.2	49.1
One NYD child	9.7	8.4	19.8	17.1	70.5	60.7
Total (70 samples)	57.5	38.9	15.6	10.7	19.7	12.6

ASD: autism spectrum disorder; DD: developmental delay; TD: typical development; NYD: not-yet-diagnosed

Table 5 Scores of the relative and absolute preferences for robots 1, 2, and 3 considering the influence of parental guidance

Sample	Relative (P_{nr}) and absolute (P_{na}) preferences for the three robots (n=1, 2, 3)					
	Robot 1		Robot 2		Robot 3	
	P_{1r}	P_{1a}	P_{2r}	P_{2a}	P_{3r}	P_{3a}
40 children without parents' guide (28 ASD and 10 DD)	56.3	41.4	17.6	13.5	25.3	18.4
22 children with parents' strong guide (15 ASD and 7 DD)	28.4	18.2	6.3	3.5	6.2	2.3
8 children with parents' weak guide (7 ASD and one DD)	47.4	23.1	11.6	6.1	7.3	2.4
Total (70 samples)	44.0	27.6	11.8	7.7	12.9	7.7

ASD: autism spectrum disorder; DD: developmental delay; TD: typical development

Table 6 Functionalities of the robots children played with most

Functionality	Frequency/ person-time
Touching sensing with a feedback	39
Operating (e.g., finding games to play or videos to watch)	31
Singing and dancing	13
Speech interacting	5
Imitating	4
Hugging	3

It is worth mentioning that most of the functionalities listed above were observed as the functionalities of robot 1 (Dabao). This highlights the importance of equipping a touching sensing with feedback functionality and an appropriate size of touch screen for robots for ASD therapy. In other words, a touching sensing functionality and a larger size of touch screen made robot 1 (Dabao) a more popular robot to interact with.

4 Conclusions and future work

4.1 Conclusions

In this preliminary clinical study, with respect to physical robot design, children with ASD and their parents were most attracted by the appearance and functionalities of robot 1 (Dabao) during HRI. It had attractive functionalities such as touching, operating, and singing and dancing. This could be one of the models for designing socially interactive robots for ASD therapy. For instance, when doing physical robot design, some aspects of appearance (e.g., with an appropriate size of touch screen to operate) and some functionalities (e.g., a touching sensing functionality to provide interactive feedback) should be taken into account. For evaluation of robots in therapy-like settings, on one hand, attention analysis using algorithms such as face detection and filtering, and estimation of the directions of gaze and head posture can be adopted to quantitatively measure the prosocial behaviors and actions performed by the children with ASD during the interventions; on the other hand, observing and calculating the time children spend on exploring/playing with the robots in video clips can be adopted to qualitatively analyze the behaviors and actions of such children.

4.2 Limitations

Certain limitations of the present study are summarized as follows:

First, in the whole time period (approximately 5 min) of the watching part of each session, it would be too difficult for some children with ASD or DD to maintain a good sitting position on the chairs for a long time. During some of the time periods, some of them would walk around, or shake their shoulders or heads, or even do some sitting gestures as if they were paralyzed, resulting in more difficulties in collecting effective samples and in using the OpenFace algorithm to analyze the attention of these samples continuously and stably.

Second, in the exploring part of each session, some parents induced strongly or weakly guided behaviors to guide their children to interact with the robots according to their preference. To balance the influence of parental guidance on children's behaviors interacting with the robots, we multiplied the preferences for the three robots by the weights we defined. The preliminary results indicated that with or without parental guidance, the preferences of children with ASD or DD for robots remain the same. However, it would be better if parental guidance could be limited in the first place. However, it will be difficult to limit such behaviors, given that parents are used to encouraging their children with ASD or DD if they lack the initiative to try new things. In future studies, parentally guided behaviors should be limited in the first place as much as possible, or at least parents should allow their children to take initiative to explore for a few minutes if limiting such behaviors was not practicable.

Third, for insights of video analysis, the watching part indicated that ASD/DD children just wanted to pay more attention to the robot in the middle. Since positioning of the robots could be a potential factor affecting the results, in future studies, to eliminate the positioning effect, counter-balancing the positions of different robots during the experiment should be considered.

Last but not the least, as the most important objective of the present study was designed originally to find out what will be the preferences of children with ASD for the appearances and functionalities of socially interactive robots, we did not try hard to recruit many TD children. We did recruit a lot of DD participants as a control group, given that there were many children with DD in the four different rehabilitation facilities/institutions where we did the study. In future studies, to find out whether children with ASD behave differently from other children when interacting with a specific functionality of a robot, more TD children should be recruited.

4.3 Future work

We indicated that functionalities such as "touching sensing" with feedback, operating (e.g., finding games to play or videos to watch), and singing and dancing were the most attractive functionalities that children with ASD would like to interact with. Future studies should emphasize on how to design HRI based on these robotic functionalities, and how to qualitatively analyze and quantitatively measure children's prosocial behaviors and actions induced by these robotic functionalities. For instance, a future study could use videos such as "The Transporters" played on the touch screen to train children with ASD to recognize facial expressions of people, and use singing and dancing as a reinforcer.

Acknowledgements

We thank the four rehabilitation facilities/institutions and their healthcare professionals for providing facilitation and assistance to our facilitators in this preliminary clinical study. We thank all the parents who took their children to participate in our study and supported us in many ways. We thank our guest students Shen-hong CHEN, Xiao-hong CAI, Hao-bo WANG, and Li-ping LI for their hard work in preparing and conducting this preliminary clinical study.

References

Baltrušaitis T, Robinson P, Morency LP, 2016. OpenFace: an open source facial behavior analysis toolkit. Proc IEEE Winter Conf on Applications of Computer Vision, p.1-10. https://doi.org/10.1109/WACV.2016.7477553

Begum M, Serna RW, Yanco HA, 2016. Are robots ready to deliver autism interventions? A comprehensive review. *Int J Soc Rob*, 8(2):157-181. https://doi.org/10.1007/s12369-016-0346-y

Clabaugh C, Becerra D, Deng E, et al., 2018. Month-long, in-home case study of a socially assistive robot for children with autism spectrum disorder. Companion of the ACM/IEEE Int Conf on Human–Robot Interaction, p.87-88. https://doi.org/10.1145/3173386.3177018

Coeckelbergh M, Pop C, Simut R, et al., 2016. A survey of expectations about the role of robots in robot-assisted therapy for children with ASD: ethical acceptability, trust, sociability, appearance, and attachment. *Sci Eng Ethics*, 22(1):47-65. https://doi.org/10.1007/s11948-015-9649-x

Costescu CA, Vanderborght B, David DO, 2014. The effects of robot-enhanced psychotherapy: a meta-analysis. *Rev Gen Psychol*, 18(2):127-136. https://doi.org/10.1037/gpr0000007

Diehl JJ, Schmitt LM, Villano M, et al., 2012. The clinical use of robots for individuals with autism spectrum disorders: a critical review. *Res Autism Spectr Disord*, 6(1):249-262. https://doi.org/10.1016/j.rasd.2011.05.006

English BA, Coates A, Howard A, 2017. Recognition of gestural behaviors expressed by humanoid robotic platforms for teaching affect recognition to children with autism—a healthy subjects pilot study. In: Kheddar A, Yoshida E, Ge SS, et al. (Eds.), Social Robotics. Springer, Cham. https://doi.org/10.1007/978-3-319-70022-9_56

Esteban PG, Baxter P, Belpaeme T, et al., 2017. How to build a supervised autonomous system for robot-enhanced therapy for children with autism spectrum disorder. *Paladyn J Behav Rob*, 8(1):18-38. https://doi.org/10.1515/pjbr-2017-0002

Greczek J, Matarić M, 2015. Encouraging user autonomy through robot-mediated intervention. Proc 10th ACM/IEEE Int Conf on Human–Robot Interaction Extended, p.189-190. https://doi.org/10.1145/2701973.2702714

Huijnen CAGJ, Lexis MAS, Jansens R, et al., 2016. Mapping robots to therapy and educational objectives for children with autism spectrum disorder. *J Autism Dev Disord*, 46(6):2100-2114. https://doi.org/10.1007/s10803-016-2740-6

Huijnen CAGJ, Lexis MAS, Jansens R, et al., 2017. How to implement robots in interventions for children with autism? A co-creation study involving people with autism, parents and professionals. *J Autism Dev Disord*, 47(10): 3079-3096. https://doi.org/10.1007/s10803-017-3235-9

Lee B, Xu J, Howard A, 2017. Does appearance matter? Validating engagement in therapy protocols with socially interactive humanoid robots. Proc IEEE Symp Series on Computational Intelligence, p.1-6. https://doi.org/10.1109/SSCI.2017.8285303

Matarić MJ, 2017. Socially assistive robotics: human augmentation versus automation. *Sci Rob*, 2(4), Article eaam5410. https://doi.org/10.1126/scirobotics.aam5410

Munir KM, Lavelle TA, Helm DT, et al., 2016. Autism: a global framework for action. http://www.wish.org.qa/summits/wish-2016/forum-reports/

Pennisi P, Tonacci A, Tartarisco G, et al., 2016. Autism and social robotics: a systematic review. *Autism Res*, 9(2): 165-183. https://doi.org/10.1002/aur.1527

Robins B, Dautenhahn K, 2014. Tactile interactions with a humanoid robot: novel play scenario implementations with children with autism. *Int J Soc Rob*, 6(3):397-415. https://doi.org/10.1007/s12369-014-0228-0

Scassellati B, Admoni H, Matarić M, 2012. Robots for use in autism research. *Ann Rev Biom Eng*, 14:275-294. https://doi.org/10.1146/annurev-bioeng-071811-150036

Simut R, van de Perre G, Costescu C, et al., 2016. Probogotchi:

a novel edutainment device as a bridge for interaction between a child with ASD and the typically developed sibling. *J Evid Based Psychot*, 16(1):91-112.

Sun X, Allison C, Matthews FE, et al., 2013. Prevalence of autism in mainland China, Hong Kong and Taiwan: a systematic review and meta-analysis. *Mol Autism*, 4(1), Article 7. https://doi.org/10.1186/2040-2392-4-7

Wainer J, Dautenhahn K, Robins B, et al., 2014. A pilot study with a novel setup for collaborative play of the humanoid robot KASPAR with children with autism. *Int J Soc Rob*, 6(1):45-65. https://doi.org/10.1007/s12369-013-0195-x

Wang S, Jiang M, Duchesne XM, et al., 2015. A typical visual saliency in autism spectrum disorder quantified through model-based eye tracking. *Neuron*, 88(3):604-616. https://doi.org/10.1016/j.neuron.2015.09.042

WUCAILU ASD Research Institute, 2017. Report on the Industry Development of Autism Education and Rehabilitation in China (II). Huaxia Publishing House, Beijing, China (in Chinese).

Zheng ZW, 2017. The Situation of the Life and the Learning of Two Million Autistic Children Became Severe. Modern Education News, Mar. 31, 2017 (in Chinese).

List of electronic supplementary materials

Fig. S1 Appearances of the three different robots: they all had a touch screen in the chest of the body, but differed in the size and the color of the body, and the dexterity of the hands

Table S1 Functionalities of the three robots

Questionnaire for Children Companion Robots