Could social robots facilitate children with autism spectrum disorders in learning distrust and deception?

Citation for the original published paper (version of record):
Zhang, Y., Song, W., Tan, Z. et al (2019). Could social robots facilitate children with autism spectrum disorders in learning distrust and
deception?. Computers in Human Behavior, 98: 140-149. http://dx.doi.org/10.1016/j.chb.2019.04.008

(article starts on next page)

Computers in Human Behavior 98 (2019) 140–149

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

Could social robots facilitate children with autism spectrum disorders in learning distrust and deception?

Yaoxin Zhang[a], Wenxu Song[b], Zhenlin Tan[b], Huilin Zhu[c], Yuyin Wang[d], Cheuk Man Lam[e], Yifang Weng[f], Sio Pan Hoi[a], Haoyang Lu[g,h], Bella Siu Man Chan[i], Jiajia Chen[b,j,**], Li Yi[a,*]

[a] School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China
[b] ZJU-SCNU Joint Research Centre of Photonics, South China Normal University, Guangzhou, China
[c] Children Developmental & Behavioral Center, Third Affiliated Hospital, Sun Yat-Sen University, Guangzhou, China
[d] Department of Psychology, Sun Yat-sen University, Guangzhou, China
[e] Institute of Psychology, Chinese Academy of Science, Beijing, China
[f] Department of Psychology and Human Biology, University of Toronto, Ontario, Canada
[g] Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China
[h] Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China
[i] Faculty of Education, The University of British Columbia, Vancouver, Canada
[j] Department of Electrical Engineering, Chalmers University of Technology, Göteborg, Sweden

ARTICLE INFO

Keywords:
Autism spectrum disorder
Social robots
Trust
Deception
Anthropomorphic thinking

ABSTRACT

Social robots have been increasingly involved in our daily lives and provide a new environment for children's growth. The current study aimed to examine how children with and without Autism Spectrum Disorders (ASD) learned complex social rules from a social robot through distrust and deception games. Twenty children with ASD between the ages of 5–8 and 20 typically-developing (TD) peers whose age and IQ were matched participated in distrust and deception tasks along with an interview about their perception of the human-likeness of the robot. The results demonstrated that: 1) children with ASD were slower to learn to and less likely to distrust and deceive a social robot than TD children and 2) children with ASD who perceived the robot to appear more human-like had more difficulty in learning to distrust the robot. Besides, by comparing to a previous study the results showed that children with ASD appeared to have more difficulty in learning to distrust a human compared to a robot, particularly in the early phase of learning. Overall, our study verified that social robots could facilitate children with ASD's learning of some social rules and showed that children's perception of the robot plays an important role in their social learning, which provides insights on robot design and its clinical applications in ASD intervention.

Our daily lives have changed dramatically thanks to innovations in information and communication technologies, which make the science fiction of the past into commonplace reality. Social robots that have a human-like appearance, also referred to as humanoid robots, are often capable of sensory motor tasks, autonomous movements, and verbal and non-verbal communications (Fong, Nourbakhsh, & Dautenhahn, 2003; Jipson & Gelman, 2007). They are increasingly able to substitute for humans in the labor force, as childminders, office receptionists, and tour guides (Itakura et al., 2008; Moriguchi, Kanda, Ishiguro, Shimada, & Itakura, 2011). It is evident that social robots will have increased participation in our daily lives such that in the future children may grow up in an environment where social robots are ubiquitous.

The growing presence of social robots makes life more convenient, but also raises certain questions, such as how children perceive and interact with robots, and how that influences their development and learning. Some research efforts have been made to address such questions. Kahn, Friedman, Pérezgranados, and Freier (2006) found that preschoolers perceive a robotic dog and a stuffed dog similarly in terms of animacy, biological properties, mental states, social rapport, and moral standing. Kahn et al. (2012) found that most children perceived that a robot had mental states (had feelings), was a social being (could be a friend), deserved fair treatment, and should not be harmed psychologically, but did not believe the robot was entitled to its own liberty or civil rights. This implied that children treated the humanoid

* Corresponding author. School of Psychological and Cognitive Sciences, Peking University, 5 Yiheyuan Road, Beijing, 100871. China.
** Corresponding author. Department of Electrical Engineering, Chalmers University of Technology, Göteborg. Sweden.
E-mail addresses: jiajiac@chalmers.se (J. Chen), yilipku@pku.edu.cn (L. Yi).

https://doi.org/10.1016/j.chb.2019.04.008
Received 8 June 2018; Received in revised form 8 April 2019; Accepted 8 April 2019
Available online 09 April 2019

robot in an anthropomorphic way to some extent, but did not perceive the robot as completely human. Thus, Kahn, Gary, and Shen (2013) proposed the new ontological category (NOC) hypothesis, that children do not categorize a social robot as either animate or inanimate, but as a new ontological entity with unique properties. In addition, when interacting with social robots, children treated them not only as social companions (Shiomi, Kanda, Ishiguro, & Hagita, 2007), but also as knowledgeable and informative interlocutors (Breazeal et al., 2016). Furthermore, children could learn knowledge and skills from robots through imitative learning, facilitated through social signals, such as eye contact (Itakura et al., 2008). Children could also learn words from a robot (Moriguchi et al., 2011).

In addition to their educational roles for typically developing (TD) children, robots have also been introduced to help children with special needs acquire knowledge and skills, especially children with Autism Spectrum Disorder (ASD), a neurodevelopmental disorder characterized by the impairments in social-communication abilities and restricted and repetitive behaviors (American Psychiatric Association, 2013). Social robots have shown many advantages in educating children with ASD. When interacting with robots, children with ASD have shown more interest (Dautenhahn & Billard, 2002, pp. 179–190), experienced more elevated attention (Scassellati, Admoni, & Matarić, 2012), and engaged more in a reversal learning task (Costescu, Vanderborght, & David, 2014), compared to interacting with people. When interacting with a humanoid robot, children with ASD are also more likely to maintain a calm and active mood (Kozima, Nakagawa, & Yasuda, 2005), show themselves to be more comfortable with emotional response modification (Ueyama, 2015), and are less likely to display repetitive behaviors (Lee, Takehashi, Nagai, Obinata, & Stefanov, 2012). These findings indicate that children with ASD might be more likely to interact with and be more comfortable interacting with robots than with people, and that such interactions can have a positive impact on their learning process. Given the specific impairments in social learning in children with ASD (Bushwick, 2001), it is important to understand the feasibility of using social robots in teaching social rules to children with ASD. The current study aimed to explore whether children with ASD would learn from robot-child interactions by creating a unique experimental setup for children with ASD to learn two important forms of social skills – distrust and deception.

The emergence of distrust and deception behaviors are two major milestones in the social-cognitive development of children that profoundly influence their real life (e.g., Lee, 2013). Other people's testimony is an important source of knowledge. To learn efficiently, children have to evaluate the reliability of the informants. However, young preschoolers, especially younger than four years of age, tend to trust people unconditionally and indiscriminately (Clément, Koenig, & Harris, 2004; Jaswal, Croft, Setia, & Cole, 2010). With increased age, preschoolers gradually acquire more complicated information and learn to process and identify it within a more expansive social milieu (e.g., Lee & Cameron, 2000), and learn to evaluate whether informants deserve trust on the basis of various factors, such as their past reliability, intentions, and past deceiving behaviors (e.g., Corriveau & Harris, 2009; Koenig & Harris, 2005; Vanderbilt, Liu, & Heyman, 2011). Not until they reach 4–5 years old do children develop skepticism towards others' testimony and selective trust, which can help them to identify reliable sources of information and foster their knowledge learning (Vanderbilt et al., 2011). In order to protect themselves from being misled by false information and to acquire credible information, children learn to distrust unreliable sources. Children also develop selective learning and the ability to perceive relativity, which assists them in coping with more advanced and complex information in their progress towards adulthood (Yi et al., 2014).

Closely related to distrust, deception behaviors also mark preschool-age social development. Deception behaviors typically begin around three years of age (Lewis, Stanger, & Sullivan, 1989), and as children grow, they deceive others with higher frequency and more sophisticated strategies (Chandler, Fritz, & Hala, 1989; Russell, Mauthner, Sharpe, & Tidswell, 1991; Sodian, 1991). At four to five years, children begin deliberately deceiving their opponents in more situations (DeVries, 1970; Gratch, 1964; Shultz & Cloghesy, 1981). Although parents may worry that their children's lying behavior may have bad consequences, a more contemporary view holds that deception may have positive effects on children's cognitive skills (Ding, Heyman, Fu, Zhu, & Lee, 2018). Evidence comes from the recent finding that learning to deceive in a four-day training program could enhance preschoolers' Theory of Mind (ToM) and executive function (Ding et al., 2018).

Typically, children learn distrust and deception rules through interpersonal interactions and based on various verbal and nonverbal social cues in past experiences (Csibra & Gergely, 2006; Lee & Cameron, 2000). It is not surprising, then, that children with ASD, who have social learning impairments, display difficulties in learning to distrust and deceive others. Yi et al. (2014) developed distrust and deception tasks in which children were asked to join in a hide-and-seek game with the experimenter always offering misleading information about the location of prizes; the experimenter and children then switched roles, and the children had the opportunity to deceive the same experimenter back about prize locations. Results indicated impairments in learning to distrust and deceive others within a social context; children with ASD showed a trust bias toward the adult who repeatedly deceived them, and were less likely to reciprocate deception. In a follow-up study, where the social components of the distrust and deception tasks were minimized (using physical rather than social cues), children with ASD showed comparable learning performance with their TD peers (Yang et al., 2017). Another study that examined trust-building in adults with ASD confirmed difficulty in social learning, particularly encoding incoming social information and applying it to update social expectations (Maurer, Chambon, Bourgeois-Gironde, Leboyer, & Zalla, 2017). Such findings indicate that individuals with ASD have difficulty learning social rules, and by reducing the social components in the learning process, children with ASD may experience less difficulty in acquiring these skills.

In the present study, we aimed to examine whether children with ASD would show similar difficulty in learning to distrust and deceive social robots as in interpersonal interactions. We used distrust and deception tasks adapted from Yi et al. (2014) and Yang et al. (2017) by replacing the role of the adult human with a humanoid robot. In the distrust task, the robot always provided misleading information about the location of hidden prizes, and children had to learn to distrust the robot in order to win the prizes. In the deception task, children needed to reciprocally deceive the robot to gain more prizes.

This study was designed to address two research questions. First, could children with ASD learn to distrust and deceive a humanoid robot from their interactions with it compared to TD children? Based on the previous research showing that social robots were perceived as a social being having mental states (e.g., Kahn et al., 2006; Kahn et al., 2012), we expected that children with ASD would still experience more difficulty in learning to distrust and deceive robots than TD children.

Second, is the learning of distrust and deception correlated with the child's perception of the robot? Previous research has found that robots are perceived as neither animate nor inanimate, but as a new category of entity (Kahn et al., 2013). However, other research suggests that both children with and without ASD perceive robots as toys, and children with ASD also perceive them as machines (e.g., Peca, Simut, Pintea, Costescu, & Vanderborght, 2014). Additionally, adolescents with ASD have been found to have more human impressions of ‘humanness’ for the robot than TD adolescents (Kumazaki et al., 2018). This discrepancy in perception of the robot might affect children's learning of distrust and deception. Breazeal et al. reported that children were sensitive to the social responsiveness of the robot, and were willing to learn from the robot with contingent responsiveness (2016). Based on this finding, we speculated that how children perceived the robot, especially their assessment of its human-likeness, would affect their performance in distrust and deception tasks. In the current study, children's anthropomorphic perception of the robot, and the role this perception played in learning distrust and deception, was assessed using questionnaires. We expected that children with ASD with more anthropomorphic thinking of the robot would experience greater difficulty in learning distrust and deception.

1. Method

1.1. Participants

Twenty Chinese children with ASD (5.08–8.83 years, $M = 6.79$, $SD = 0.93$, two females) and 20 age- and IQ-matched Chinese TD children (5.24–7.32 years, $M = 6.35$, $SD = 0.56$, three females) were recruited for this study. TD children were recruited from the community and an elementary school in Beijing. Children with ASD were recruited from two Chinese cities: Beijing (14 males and two females, tested before intervention) and Dongguan (four males). All children with ASD in our study were previously diagnosed by experienced hospital pediatricians based on the DSM-IV-TR diagnostic criteria (American Psychiatric Association, 2000, pp. 553–557). Additionally, the diagnoses of 16 out of 20 children with ASD in Beijing were further confirmed by the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) and Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, & Le Couteur, 1994). Among these children, 15 scored between 6 and 10 on total severity, falling in the "typical autism" range and one scored 5, falling in the ASD range. (Gotham, Pickles, & Lord, 2009; see Table 1 for details). The remaining four children with ASD, recruited in Dongguan where no reliable examiner for ADOS and ADIR was available, were assessed using the Social Responsive Scale (SRS; Constantino & Gruber, 2005; all above the cut-off scores) based on parental reports. TD children were recruited from typical primary schools in Beijing to match the ASD group according to chronological age and IQ, measured by Wechsler Intelligence Scale for Children - Fourth Edition (WISC-IV; Wechsler, 2003; see Table 1).

To further compare if children with ASD could distrust and deceive a robot versus a real person, we also included another dataset from the previously published study using the similar distrust and deception tasks involving interpersonal interactions (see social condition in Yang et al., 2017). This dataset included 17 TD children (mean age = 5.72, $SD = 0.69$, age range 4.00–7.00 years) and 22 children with ASD (mean age = 6.09, $SD = 0.76$, age range 4.62–7.62 years). IQs were measured by Combined Raven Test (CRT-C2). For the TD group the average of CRT standard score was 100.69 ($SD = 13.51$), and for the ASD group the average score was 100.59 ($SD = 16.25$). However, the mean age in Yang et al. (2017)'s study was significantly younger than the current sample, $t(77) = 2.89$, $p = .005$. This could confound the findings, as group differences could be attributed to age difference instead of difference between the human and robot conditions. Therefore, for the ASD group, given the fact that it was difficult to recruit more ASD participants, we decided to remove the youngest children with ASD one by one until the ages of these two ASD groups could be matched (we found that 4 is the minimum number of children to be removed to match these two groups). We also excluded one child with ASD due to missing trial by trial data. Then this group of 17 children with ASD (mean age = 6.31, $SD = 0.65$) could be matched with our ASD group in the robot condition, $t(35) = -1.81$, $p = .079$ (see Fig. 1). For the TD group, as recruiting TD children was feasible, we reran the experiments of Yang et al. (2017) with 5 extra TD children, all older than the mean age of Yang et al. (2017)'s TD sample, by strictly following the same procedure. Together with the original Yang et al. (2017)'s 17 TD children, this group of TD children (mean age = 6.17, $SD = 0.77$) could be matched with our TD group in the robot condition, $t(40) = 0.87$, $p = .390$ (see Fig. 1). Finally, the final dataset of the human condition including 17 children with ASD (three females, mean age = 6.17, $SD = 0.77$, age range 5.00–7.92 years; Mean IQ = 100.48, $SD = 12.03$) and 22 TD children (two females, mean age = 6.31, $SD = 0.65$, age range 5.43–7.62 years; Mean IQ = 99.88, $SD = 14.18$) could be matched with each other, $t(37) = -0.58$, $p = .564$ (see Fig. 1 for details). We conducted an independent sample Welch t-test to compare age and IQ differences between these four groups (robot-ASD group, robot-TD group, human-ASD group and human-TD group); there was no significant difference, $ps > .079$ (Bonferroni corrected). Also, to make sure that the symptom severity was similar between the ASD-human group and the ASD-robot group, we compared their AQ and SRS scores using the independent sample Welch t tests, and found a non-significant group difference in the symptom severity, for AQ: $t(35) = -5.42$, $p = .591$, for SRS: $t(35) = -1.12$, $p = .272$.

Table 1
Characteristics of the participants in each group.

Characteristics	ASD	TD	Group Differences
N	20	20	N/A
Male (female)	18 (2)	17 (3)	N/A
Age range (years)	5.08–8.83	5.24–7.32	N/A
Age in years (SD)	6.79 (0.93)	6.35 (0.56)	$t(31) = 1.80$, $p = .081$
IQ (SD)	104.90 (11.90)	106.70 (12.63)	$t(38) = -0.46$, $p = .645$
AQ ($N = 20$)	81.40(17.78)	N/A	N/A
SRS ($N = 20$)	82.85 (24.32)	N/A	N/A
ADOS Total ($N = 16$)	17.19 (4.67)	N/A	N/A
Social–Communication	13.63 (3.59)	N/A	N/A
RRB	2.88 (1.02)	N/A	N/A
ADOS Total Severity ($N = 16$)	8.19 (1.76)	N/A	N/A
SA Severity	8.19 (1.47)	N/A	N/A
RRB Severity	7.50 (1.37)	N/A	N/A
ADI-R Total ($N = 16$)	51.06 (12.99)	N/A	N/A
Social–Communication	39.50 (11.23)	N/A	N/A
RRB	8.50 (2.37)	N/A	N/A

Note. SA = Social Affect; RRB = Restricted, Repetitive Behavior (Gotham et al., 2008).

1.2. Social robot

We used the robot *Nao* (developed by Aldebaran Robotics, France, see Fig. 2) to play the distrust and deception games with the children. *Nao* is a 58-cm high and 5-kg humanoid robot with 25 degrees of freedom. *Nao* moves agilely, with an inertial navigation device to maintain stability, and can detect and bypass obstructions using two pairs of ultrasonic transmitters and receivers, which allow for accurate motion. *Nao* is balanced by four pressure sensors controlling the corresponding pressure center on each foot. It has four loudspeakers and one speech recognition and analysis system, which allows it to listen, speak, and perform space-sound positioning, and two high-definition CMOS cameras that enable forward vision; such powerful hardware gives *Nao* a high degree of artificial intelligence. In this study, *Nao* communicated with the children through behaviors including walking, swinging, pointing, blinking, speaking and so forth. The experiment was designed to conform to ecologically valid concepts. *Nao* has been used previously with children with ASD and appears to be attractive to them (Boucenna, Anzalone, Tilmont, Cohen, & Chetouani, 2014). *Nao* was called "*Naonao*" in our study to resemble the pronunciation of a child's name in Chinese.

1.3. Procedure

In the experiment, lasting approximately 25 min, each child participated in a series of tasks in the following order: warm-up session, distrust and deception tasks, and a short interview about their anthropomorphic thinking of the robot. The first author of the paper, who

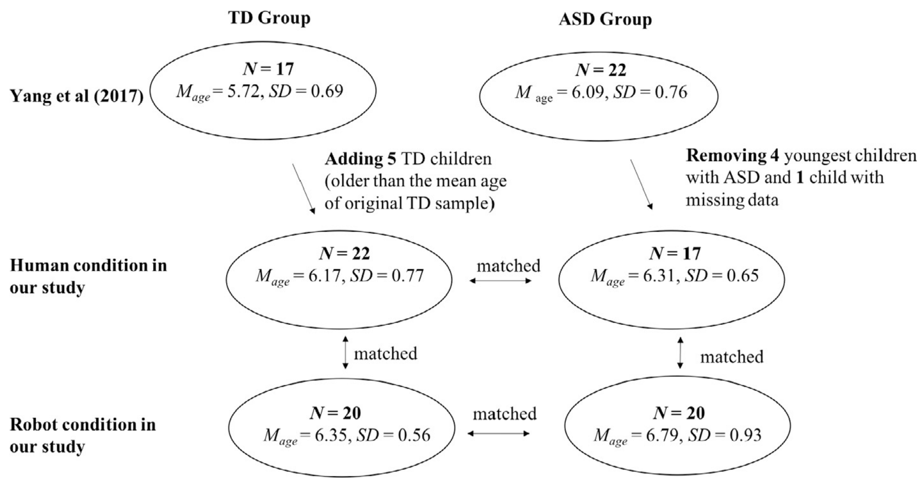

Fig. 1. The illustration for the sources of the participants.

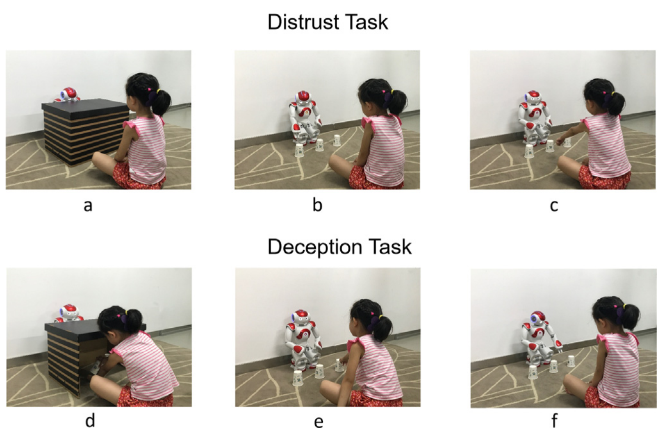

Fig. 2. The procedure of distrust (a–c) and deception task (e–f). In the distrust task, after the experimenter hid two tokens under two of the three cups, the robot pretended to check all three cups under the big carton (a). Then the carton was removed, and the robot showed the child the location of the token by pointing to one cup (b). The child was then asked to choose only one cup that they thought contained the token by pointing (c). In the deception task, the experimenter used the carton to cover three cups when the child was hiding the token under the cups (e). Then the carton was removed and the child were asked to point to one cup for the robot (f). The robot pointed to the cup indicated by the child and said "I guess the token is here" (g).

was trained beforehand on performing the tasks, acted as the experimenter for the ASD and the TD groups to control the potential impact of different experimenters.

Warm-up session. The experimenter introduced the robot *Naonao* to children, and invited them to engage in a brief (two-to-three minute) semi-structured warm-up conversation with the robot. The children were encouraged to chat with the robot by asking it questions (e.g., "what is your name?", "how old are you?", "what do you like to eat?", and "what is your favorite fruit?"). *Naonao* answered these questions and asked similar questions back to children. This warm-up session was designed to help participants get familiar with *Naonao* and relieve any potential tension when facing a robot.

Distrust and deception tasks. A child and *Naonao* sat on a mat face-to-face, and three identical upside-down opaque cups were placed between them. One female experimenter sat to the side of *Naonao* and the child, facing the three cups. At the task onset, the experimenter invited the child to participate in a hide-and-seek game with the robot to search for tokens. All participants were told that the more tokens they won, the more prizes they would get upon completion of the experiment. The procedures of the distrust and the deception tasks are illustrated in Fig. 2.

In the distrust task (Fig. 2, a-d), the experimenter first introduced the rules of the task: If the children successfully found the token, they could keep it; otherwise, *Naonao* would get it. The experimenter placed a modified paperboard carton (two of six faces removed, see Fig. 2) over the cups to prevent the children from peeking during the hiding process. The experimenter then hid two tokens under two of the three cups, but told the children only one cup contained the token and they had to find it with *Naonao*'s help. *Naonao* would pretend to check all three cups under the carton (Fig. 2a). The carton was removed, and *Naonao* showed the children the token's alleged location by pointing at one cup, saying "the token is here" (Fig. 2b). Note that *Naonao* was pre-programmed to always point at the empty cup, while the other two cups contained the tokens. The children were asked to choose the cup they thought contained the token by pointing (Fig. 2c). After the children's decision, the experimenter lifted the upside-down cup to check whether the children had found the token. The experimenter then informed the children whether they had won the token. This hide-and-seek procedure was repeated ten times.

For the deception task (Fig. 2, e-f), children were asked to hide the token for *Naonao* to seek. The experimenter used the carton to cover three cups while children hid a token under one of the cups (Fig. 2e). The carton was removed, and children were asked to point to one cup for the robot (Fig. 2f). *Naonao* then pointed to the cup indicated by children and said "I guess the token is here" (Fig. 2g). For simplicity, *Naonao* was programmed to always trust the children. The experimenter would lift the cup *Naonao* selected, and provide feedback accordingly. This process was also repeated ten times.

For each task, if children successfully distrusted or deceived *Naonao* in five consecutive trials, they were considered to have successfully passed the distrust or deception task and the remaining trials were marked as correct. In each trial, participants scored 0 when they trusted or did not (incorrect) deceive *Naonao*, or scored 1 when they distrusted or deceived the robot (correct). Performance scores were averaged across ten trials.

Anthropomorphic thinking questions. Following the distrust and the deception tasks, we asked participants several yes-or-no questions in order to investigate their perception of the robot, especially their anthropomorphic thinking. These questions were adapted from previous literature (Kahn et al., 2006; Martini, Gonzalez, & Wiese, 2016; Rusca & Tonucci, 1992) to reflect three aspects of the children's conceptions of robotic entities: animacy and other biological properties (questions a-e), mental states (questions f-g), and social rapport (question h). Questions were listed as follows:

(a) Is *Naonao* alive?

(b) Does *Naonao* eat?
(c) Does *Naonao* sleep?
(d) Will *Naonao* grow up?
(e) Does *Naonao* have parents?
(f) If you beat *Naonao*, would it[1] feel pain?
(g) If you give gifts to *Naonao*, would it feel happy?
(h) Do you like *Naonao*?

1.4. Data analysis

We used Welch t tests to compare group differences in total scores across all ten trials of the distrust and the deception tasks respectively. Participants learned from trial-by-trial feedback (whether they gained a token or not) in the tasks; to investigate this learning process, we further conducted survival analyses as well as log-rank tests to examine how distrust and deception learning curves differed between groups. The survival analyses quantified the percentage of children who stably learned to distrust or deceive the robot (gaining tokens in five successive trials), and log-rank tests compared learning speeds between the ASD and TD groups. Then, each child's distrust and deception performances were correlated with their verbal answers to the interview regarding anthropomorphic thinking. Lastly, the performance in the robot condition were compared with the previously published data of the performances in interpersonal interactions (Yang et al., 2017) to examine the facilitation effect of the human-robot interaction on distrust and deception learning.

2. Results

2.1. Distrust task

The independent-sample Welch t-test showed a significant difference in overall distrust performance between the ASD ($M = 7.70$, $SD = 2.62$) and TD groups ($M = 9.35$, $SD = 0.67$), $t(21.49) = -2.73$, $p = .012$. TD children were more likely than children with ASD to distrust the robot who offered incorrect information.

Survival analyses were conducted to measure children's learning process throughout all 10 trials in the distrust task. As shown in Fig. 3a, the curves showed the average percentage of children who failed to distrust at different time points, and the shaded areas represent the 95% confidence intervals. Results indicated that the TD group demonstrated a sharp downward trend: 85% began to distrust before the third trial, and all TD children successfully passed the task after 10 trials. However, the learning curve of children with ASD was less steep and only 70% passed the task after 10 trials. In order to examine the differences in the speed of learning to distrust the robot between children with ASD and TD, we conducted log-rank tests between the two groups, and the result showed a significant group difference in *learning time costs* (defined as the number of trials that reflect the learning time needed to pass the task), χ^2 (1) = 8.03, $p < .005$. That is, the ASD group learned to distrust the robot less efficiently than the TD group.

2.2. Deception task

The overall performance analysis of the deception task also found a significant difference of the overall deception performance between the ASD ($M = 6.70$, $SD = 3.64$) and the TD ($M = 9.55$, $SD = 1.19$) groups, $t(23.02) = -3.33$, $p = .003$, indicating that TD children were more likely to deceive the robot than children with ASD.

We also conducted a similar survival analysis for the deception performance (see Fig. 3b). Similar to the distrust task, the TD group showed a steep downward trend: 90% began to deceive before the third trial, and after all 10 trials, 90% passed the task (successfully deceived *Naonao* in five consecutive trials). However, the learning curve of the ASD group was less steep, and only 55% successfully passed the task after all 10 trials. The log-rank tests showed a significant group difference in *learning time costs*, χ^2 (1) = 6.93, $p = .008$, suggesting that the ASD group was significantly slower than the TD group in learning to deceive the robot.

2.3. Anthropomorphic thinking and its correlation with distrust and deception performance

For the eight questions about their anthropomorphic thinking of the robot, participants' affirmative answers ("yes") indicated the perceived human-likeness of the robot. The percentage of affirmative responses in ASD and TD groups and group difference are listed in Table 2; internal consistency of all questions was 0.78 (Cronbach's alpha value).

We conducted chi-square tests to examine whether the ASD and TD groups differed in their anthropomorphic thinking of the robot, and found that the two groups responded very similarly except on two questions: The ASD group was more likely than the TD group to think that the robot could grow up, χ^2 (1) = 4.80, $p = .028$, and marginally more likely to think that the robot would feel pain, χ^2 (1) = 3.60, $p = .058$ (see Table 2). We further summed up the scores of all eight questions and compared the total scores between the groups. An independent Welch t-test found a marginal group difference, t (38) = 1.84, $p = .074$: the ASD group ($M = 4.35$, $SD = 2.32$) attributed slightly more anthropomorphic thinking to the robot than the TD group ($M = 3.10$, $SD = 1.97$). These results suggest that although both groups did not consider the robot human-like in every respect, children with ASD seemed to be more prone to attribute anthropomorphic thinking to the robot than their TD peers.

More importantly, we tested the correlation of participants' anthropomorphic thinking with overall performance in the distrust and deception tasks by using the Pearson correlations coefficients. First, we conducted the correlations between all participants' (both groups combined) anthropomorphic thinking of the robot and distrust/deception performance, and found that participants' anthropomorphic thinking were negatively correlated with distrust performance, r (38) = −0.46, $p = .003$, but not with deception performance, r (38) = −0.25, $p = .115$. Children who attributed more anthropomorphic thinking to the robot were less likely to distrust the robot. Second, we conducted correlations between anthropomorphic thinking and distrust/deception performances for each group separately (see Fig. 4). The results showed that this negative correlation between anthropomorphic thinking and distrust performance only existed for the ASD group, $r(18) = -0.46$, $p = .042$, suggesting that children with ASD who attributed more anthropomorphic thinking to the robot were less likely to learn to distrust the robot. For the TD group, no correlation was found between anthropomorphic thinking and distrust performance, $r(18) = -0.31$, $p = .189$. No correlation between anthropomorphic thinking and deception performance was found for either group, ASD: $r(18) = -0.11$, $p = .640$, TD: r (18) = −0.29, $p = .209$.

2.4. Human-robot vs. interpersonal interactions

To examine whether the human-robot interactions could facilitate the learning of social rules for children with ASD, we further compared the results of the current study with our previously published data (Yang et al., 2017) using humans rather than robots. We first conducted a 2 (Group: ASD vs. TD) × 2 (Condition: robot vs. human) ANOVA of the overall distrust performance to compare how children distrusted the robot/person. We found that a significant main effect of Group, $F(1, 75) = 37.64$, $p < .001$, $\eta_p^2 = 0.33$, a marginally significant main effect of Condition, $F(1, 75) = 3.39$, $p = .070$, $\eta_p^2 = 0.04$, and a marginally significant Group × Condition interaction, $F(1, 75) = 3.96$, $p = .050$, $\eta_p^2 = 0.05$. Further, we conducted two simple effect analyses, and

[1] In Chinese, *it* ("ta1") sounds exactly the same as *he* or *she*, so that the use of pronoun would not affect children's perception of human-likeness of the robot.

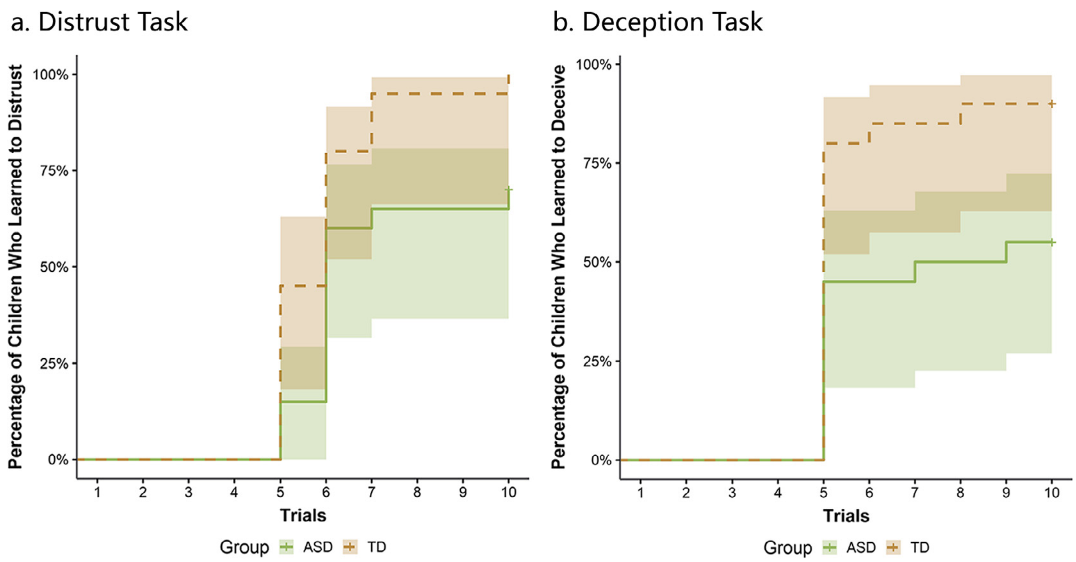

Fig. 3. Survival analysis. Percentages of children who failed to learn to distrust over trials in (a) the distrust task and (b) the deception task for ASD and TD groups in robot condition.

Table 2
Percentage of children who gave affirmative answers to the questions about their perception of the robot.

Question Categories	Question	TD	ASD	Group Differences
Animacy and Other Biological Properties	a. Is Robot *Naonao* alive?	35%	55%	χ^2 (1) = 1.62, p = .204
	b. Does *Naonao* eat?	10%	25%	χ^2 (1) = 1.56, p = .212
	c. Does *Naonao* sleep?	35%	50%	χ^2 (1) = 0.92, p = .337
	d. Would *Naonao* grow up?	10%	40%	χ^2 (1) = 4.80, p = .028
	e. Does *Naonao* have parents?	75%	75%	χ^2 (1) = 0, p = 1
Mental States	f. If you beat *Naonao*, would it feel pain?	35%	65%	χ^2 (1) = 3.60, p = .058
	g. If you give gifts to *Naonao*, would it feel happy?	15%	35%	χ^2 (1) = 2.13, p = .144
Social Rapport	h. Do you like *Naonao*?	95%	90%	χ^2 (1) = 0.36, p = .548

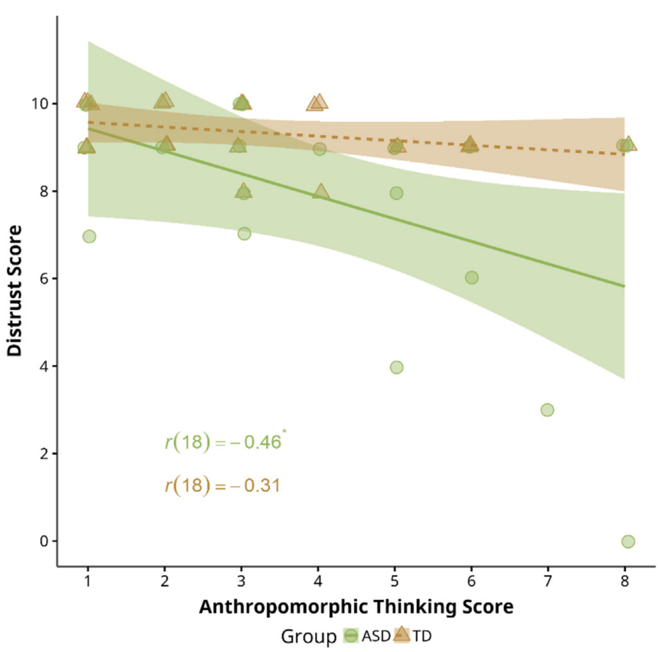

Fig. 4. Correlations between the anthropomorphic thinking of the robot and the total scores of the distrust task in the ASD and the TD groups.

found that the children with ASD were more likely to learn to distrust the robot (M = 7.70, SD = 2.62) than the person (M = 6.18, SD = 2.30), $F(1, 75)$ = 6.88, p = .011, η_p^2 = 0.08; while TD children showed similar distrust performance whether facing the person (M = 9.41, SD = 0.67) or the robot (M = 9.35, SD = 0.67), $F(1, 75)$ = 0.01, p = .914, $\eta_p^2 < 0.01$ (see Fig. 5a). Also, although the group difference for the robot and the human conditions were both significant, the group effect of distrust performances in the human condition, $F(1, 75)$ = 32.32, $p < .001$, η_p^2 = 0.30, was greater than the robot condition, $F(1, 75)$ = 8.78, p = .004, η_p^2 = 0.11. This finding implies that the learning of distrust was higher when children with ASD interacted with the robot.

To examine when this facilitation effect of the robot occurred, we further conducted a temporal analysis by dividing the 10 trials into early (Trials 1–3), middle (Trials 4–7), and late (Trials 8–10) phases, and compared distrust performance between groups and conditions for each phase separately. In the early phase, we found a significant main effect of Group, $F(1, 75)$ = 35.78, $p < .001$, η_p^2 = 0.32, and a Group × Condition interaction, $F(1, 75)$ = 5.06, p = .027, η_p^2 = 0.06. Simple effect analyses showed that TD children performed similarly in robot and human conditions, $F(1, 75)$ = 0.21, p = .648 $\eta_p^2 < 0.01$, whereas children with ASD were more likely to distrust the robot than the person, $F(1, 75)$ = 7.06, p = .010, η_p^2 = 0.09 (see Table 3). Similar main effects of the group were found in distrust performance in the middle and late phases, $ps < .001$; however, neither Group × Condition interaction or the Condition effect was found for these phases, $ps > .115$ (see Table 3 for more details). These findings indicate that the facilitation effect of the robot for distrust performance in children with ASD occurred at the early phase of learning, and disappeared in the middle and late phases.

For the deception task, we performed a similar Group (ASD vs. TD) × Condition (robot vs. human) ANOVA to compare overall deception performance in the robot and human conditions. As shown in

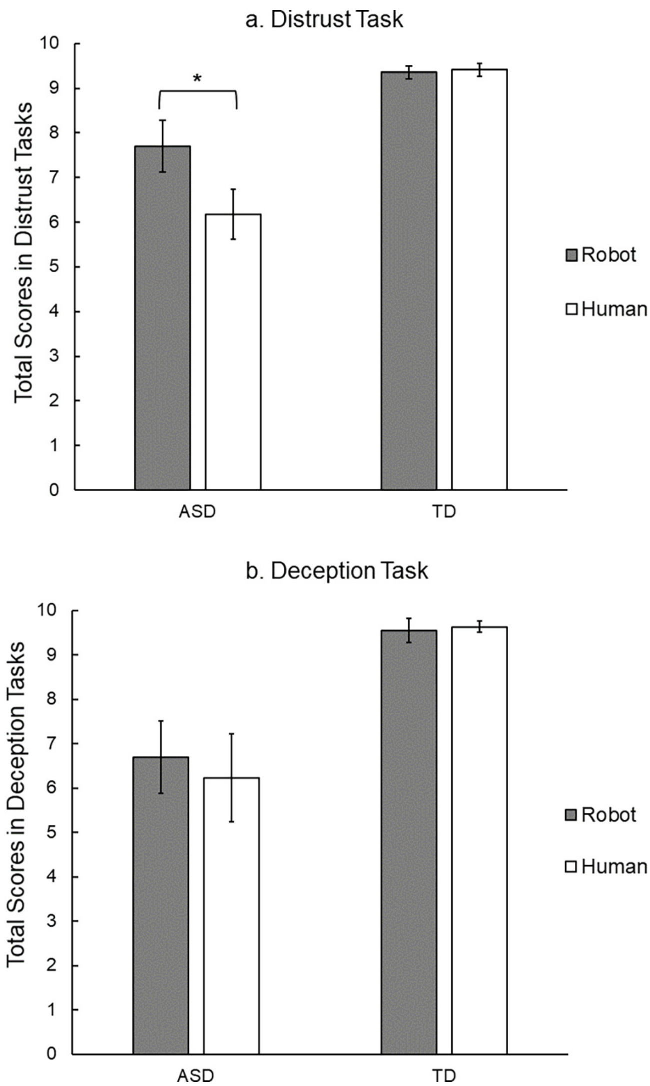

Fig. 5. Total distrust and deception scores of the ASD and the TD groups in the robot and the human conditions.

Fig. 5b, only the main effect of Group was significant, $F(1, 75) = 26.02$, $p < .001$, $\eta_p^2 = 0.26$. Neither the main effect of Condition nor the Group × Condition interaction was significant, $ps > .65$ (see Table 3). In both conditions, the ASD group was less likely to deceive the agent than the TD group.

We also performed a similar temporal analysis for deception performance by conducting the Group (ASD vs. TD) × Condition (robot vs. human) ANOVAs to compare how ASD and TD children deceive the robot and the person at different phases. Results, as listed in Table 3, showed the significant main effects of Group in all of the three phases, $ps < .001$, but no significant main effects of Condition or the Group × Condition interaction, $ps > .307$. These findings suggest that children with ASD were less likely to learn to deceive both agents than TD children across all time phases.

3. Discussion

As social robots gradually step into modern society, we must explore the various conveniences they can bring. There is a growing body of literature describing how children and people with special needs can benefit from interacting with and learning from social robots (e.g., Dautenhahn & Billard, 2002, pp. 179–190; Palsbo & Hood-Szivek, 2012; Scassellati et al., 2012). In the current study, we examined whether children with ASD would demonstrate difficulty in learning social rules (e.g., distrust and deception) when interacting with social robots, and we investigated whether this learning of social rules from robots would depend on children's perception of the robot, especially the perceived human-likeness of the robot. Particularly, we implemented the distrust and deception games to establish experimental situations and provided findings in terms of the two aforementioned aspects. First, the results showed that children with ASD were slower and less likely to learn to distrust and deceive a social robot than their TD peers, suggesting that children with ASD still experience difficulty in learning complex social rules when interacting with a social robot. Second, we investigated the impact of anthropomorphic thinking on learning distrust and deception, and found that children with ASD, who perceived the robot to be more human-like, experienced more difficulty in learning to distrust the robot.

Furthermore, we examined whether interaction with the robot facilitate social learning to be more correctly and quickly in children with ASD compared to their learning from interpersonal interactions. For this purpose, we further compared the distrust and deception performance measured in the current study with the previous ones from Yang et al. (2017), which used nearly identical tasks but with interpersonal rather than human-robot interactions to examine whether a robot would enhance the learning of social rules for children with ASD. We found that children with ASD performed better in learning to distrust a robot than a real person, particularly in the early phase of the distrust task.

3.1. Distrust and deception learning in ASD

Previous studies have shown that children with ASD are less likely to learn to distrust people who repeatedly deceived them and less efficient in doing so, and also are less likely to deceive back (Yang et al., 2017; Yi et al., 2014). Our study found similar patterns in children with ASD when interacting with a humanoid robot: They were also less likely to learn to distrust and deceive a social robot compared to TD children and less efficient in doing so. In our study, we used a social robot to teach children with ASD complex social rules, and attested that teaching children with ASD was more difficult than teaching TD children.

Passing the distrust and deception tasks requires understanding and manipulation of others' mental states (Lee & Cameron, 2000; Leekam & Prior, 1994). Some researchers have reported that TD children's distrust and deception behaviors are significantly related to their development of ToM and inhibitory control (e.g., Carlson, Moses, & Hix, 1998). Note that in this study, TD children could learn to distrust and deceive, suggesting that such complex social rules can be acquired from interacting with social robots. However, our findings clearly showed that children with ASD could not perform as well as TD children in both the distrust and deception tasks. Such results might be explained by the deficient nature of ToM and inhibitory control in children with ASD, which may cause difficulty in attributing the mental states of social agents.

It is noteworthy that although children with ASD experienced more difficulty compared to TD children in learning to distrust and deceive the robot, they had a similar learning trend within the 10 trials as the TD group. The two groups did adjust their distrust and deception behaviors according to the trial-by-trial feedback, but with different progresses. This result implies that children with ASD are largely impaired but not totally deprived in all aspects of social learning, so that they could still learn to some extent from social interactions. However, due to their social impairments, their social learning is not as efficient as that of TD children.

3.2. Anthropomorphic thinking of robot and distrust

We also examined how children's distrust and deception

Table 3
The effects of group and condition on Children's distrust and deception performance at early, middle, and late phases.

	Mean (*SD*)				Effects			
			Robot	Human		Group	Condition	Group × Condition
Distrust Task	All Trials	ASD	7.70 (2.62)	6.18 (2.30)	*F*	37.64	3.39	3.96
		TD	9.35 (0.67)	9.41 (0.67)	*p*	< .001	.070	.050
					η_p^2	.33	.04	.05
	Early Phase (Trial 1–3)	ASD	1.80 (0.77)	1.18 (0.81)	*F*	35.78	2.65	5.06
		TD	2.40 (0.68)	2.50 (0.60)	*p*	< .001	.108	.027
					η_p^2	.32	.03	.06
	Middle Phase (Trial 4–7)	ASD	3.30 (1.17)	2.76 (1.09)	*F*	24.61	2.54	1.87
		TD	3.95 (0.22)	3.91 (0.29)	*p*	< .001	.115	.176
					η_p^2	.25	.03	.02
	Late Phase (Trial 8–10)	ASD	2.60 (0.88)	2.24 (1.09)	*F*	14.71	1.44	1.44
		TD	3.00 (0)	3.00 (0)	*p*	< .001	.234	.234
					η_p^2	.16	.02	.02
Deception Task	All Trials	ASD	6.70 (3.64)	6.24 (4.07)	*F*	26.02	0.10	0.20
		TD	9.55 (1.19)	9.64 (0.58)	*p*	< .001	.758	.654
					η_p^2	.26	.001	.003
	Early Phase (Trial 1–3)	ASD	1.85 (1.18)	1.53 (1.42)	*F*	22.31	0.80	0.34
		TD	2.75 (0.55)	2.68 (0.48)	*p*	< .001	.374	.563
					η_p^2	.23	.01	.004
	Middle Phase (Trial 4–7)	ASD	2.85 (1.42)	2.47 (1.77)	*F*	21.98	0.19	1.06
		TD	3.80 (0.70)	3.95 (0.21)	*p*	< .001	.666	.307
					η_p^2	.23	.002	.004
	Late Phase (Trial 8–10)	ASD	2.00 (1.21)	2.24 (1.20)	*F*	22.39	0.40	0.40
		TD	3.00 (0)	3.00 (0)	*p*	< .001	.530	.530
					η_p^2	.23	.01	.01

performance related to their anthropomorphic thinking concerning the robot, which may affect children's attribution of mental states to the robot. Some previous literature on anthropomorphic thinking in TD children has indicated that children tend to attribute anthropomorphic thinking to non-living things. For example, Kahn et al. (2012) reported that most children perceived that the robot had mental states, was a social being, deserved fair treatment, and should not be harmed psychologically. What is more, since social robots have been designed to resemble humans in many aspects (motions, language, appearance, etc.), they have been treated by children as a new species between living and non-living things, in line with the new ontological category theory (Kahn et al., 2013). Therefore, it would be reasonable to infer that children may attribute a certain level of anthropomorphic thinking to the social robots. Kumazaki et al. (2018) found that adolescents with ASD had more impressions of 'humanness' for the robot than TD adolescents. Our study confirmed this finding by showing that children with ASD tended to perceive the social robot as more human-like than TD children do. For example, they were more likely than TD children to think that the robot could grow up and experience pain. More importantly, the results of our study further attested that the perceived human-likeness of the robots might influence children's learning of distrust: Children with ASD who attributed more anthropomorphic thinking to the robot were less likely to learn to distrust the robot. We speculated that this is because children with ASD who attribute higher degrees of perceived human-likeness to the robot might generalize their social deficiency in interpersonal interaction to their interaction with the robot, thus experiencing more difficulty in understanding the deceptive intention of the robot in the distrust task. On the other hand, those with less anthropomorphic thinking could treat the robot as a toy and not need to infer its mental states, thus not generalizing their social deficiency to the human-robot interactions. It is consistent with the findings of Yang et al. (2017) that children with ASD did not have difficulty learning from physical cues, but had specific difficulty in learning from social cues.

Some previous literature has demonstrated that the characteristics of social robots, such as their appearance, manner of movement and expression, influence the way children interact with them (e.g., Breazeal, 2002). For instance, Breazeal et al. (2016) asked children to interact with contingent and non-contingent robots, and found that young children were sensitive to the social responsiveness of the robots and preferred the contingent robot as an informant. Beyond these findings, our study further indicates that children with ASD's perception of the human-likeness of robots also influences their interactions with the robot. This finding implies that how children with ASD interact with the robot will be shaped not only through manipulating the robot's characteristics, but also through simply manipulating the child's perception of it. Future research in this area could attempt to make social learning for children with ASD more efficient by adjusting their anthropomorphic perception. We expect that changing the robot's description to reduce its perceived human-likeness may potentially change the children's perception and enhance their social learning, such as emphasizing that it is actually controlled by an experimenter, but not having its own feeling and thinking. Developing various programs that enable social robots with different behavioral characteristics might change the children's perception of robots as well, boosting the learning process of some social rules.

3.3. Human-robot vs. interpersonal interactions

Despite the clear difficulty shown in learning complex social rules, our findings show that children with ASD do obviously better in interacting with a robot than with a real person in distrust tasks, as they were more likely to learn to distrust the robot, especially during the early phase of the trials. This advantage diminished over time, which we speculate was due to a change in the child's perception of the robot, specifically in increased perceived human-likeness. Children with ASD might be attracted to the robot as a new species, particularly in the beginning of the learning process. As they get more familiar with the robot by interacting with it, they gradually develop an anthropomorphic thinking of the robot, perceiving it to be more and more human-like. As discussed earlier, this development of anthropomorphic thinking could hamper the learning of distrust for children with ASD; their learning might become inefficient due to their deficiency in inferring people's minds. On the other hand, such a facilitation effect was not found in the deception task, which could be due to two factors. First, deception, which requires not only the ability to understand but also to manipulate others' belief, is more difficult than distrust in general. Second, as the deception task always followed the distrust task by design, children had already developed their anthropomorphic thinking before they participated in the deception task, which could have hampered their ability to complete the task. In summary, the findings imply potential advantages of humanoid robots in social rules and skills training for children with ASD, as robots may be beneficial not only in replacing or reducing the professional personnel required in ASD intervention, but also in facilitating the learning process for social rules such as distrust. This presents a positive sign for children with ASD in the future, in which more social robots will exist in daily life.

3.4. Limitations

Several considerations emerge from the current findings. First, it is noteworthy that this study was not exactly an intervention study to test the efficacy of robots in the ASD intervention. Instead, it should be considered a proof-of-concept study providing preliminary findings for the efficacy of a robot in ASD intervention. The “learning” in our experiment was restricted to the experimental setting, which may not generate a long-term effect and hence cannot be generalized to real-life situations. Follow-up research for testing the effects of a robot on ASD symptoms should use the randomized-controlled-trial design and multiple training programs lasting for several weeks or months, and also include pre- and post-measures of the outcome. Second, we did not use a within-subject design to compare the same participants with human and robot conditions, which would examine the potential impacts of these two conditions on each other. If a child participates in both conditions (for instance, human condition first and robot condition later), their learning in the second session would not be considered “learning from the robot” but probably a carry-over effect from the first session with a similar task. Therefore, a between-subject comparison is appropriate for this purpose. However, due to limited access to children with ASD, we have to compare the findings of this paper with the previous literature. Third, although there is no significant difference of the mean ages between the two ASD groups, the age difference of 6 months could still represent different neurodevelopmental patterns in children, which could affect their interactions with robots and humans. To address this limitation, future studies should recruit a large group of children and randomly assign them into the robot or human conditions to make sure the two groups are perfectly matched by age. Fourth, we would assume that our findings are mainly cultural-general, based on the previous literature. For example, children from both Eastern and Western societies behave very similarly in evaluating deceptive behaviors, considering both cultures' encouragement of honesty (e.g., Fu, Xu, Cameron, Heyman, & Lee, 2007; Xu, Bao, Fu, Talwar, & Lee, 2010). However, Eastern and Western children could demonstrate subtle difference in their distrust and deception development, given cultural differences in values and meanings (Rothbaum, Pott, Azuma, Miyake, & Weisz, 2000; Xu et al., 2010). Also, children from different countries may have different perceptions on robots due to different levels of exposure to them. Therefore, it is highly recommended that future research examines whether there is a cultural difference in children's interaction with robots, and how it could affect their distrust and deception learning. Fifth, since we did not investigate anthropomorphic thinking for the human condition, it is not clear whether the interaction progress would affect children’ anthropomorphic thinking answering. Thus, future research could add this point into the experimental design. Sixth, we recruited four children from Dongguan since we were not able to recruit enough participants in Beijing. Including more participants would help increase the reliability of data analysis. Since there are only a handful of reliable administrators of ADOS and ADIR in China, these four children were not able to be assessed in these two scales in Dongguan. Future research should consider the regional difference of the performance of this task, and to replicate these findings in more cities and counties. Last, we could not compare anthropomorphic thinking between robot and human interaction groups since these questions were not included in Yang et al. (2017). Future investigations on whether children would have different anthropomorphic thinking for the robot and human.

4. Conclusions

Overall, our study contributes several promising preliminary findings on the potential involvement of humanoid robots in social rules training for children with ASD. Our results also shed light for the direction of future research, which should address whether social learning from robots can be generalized to a universal case (e.g., whether distrusting/deceiving the robot contributes to an equivalent effect on distrusting/deceiving a real person); a validation test would be required in future work to test whether children with ASD who manage to distrust and deceive a robot are capable of doing the same to a real person.

Conflicts of interest

The authors report no biomedical financial interests or potential conflicts of interest.

Acknowledgements

This study was supported by National Natural Science Foundation of China (31871116, 31571135), Beijing Municipal Science and Technology Commission (Z171100000117015), Göran Gustafsson Foundation in Sweden and STINT Joint China-Sweden Mobility Programme. The authors wish to thank Shen Wang, Dandan Wang, Jiayun Guo, Ci Song, Tianbi Li, Qiandong Wang, Sixiao Zheng, Yuchen Yao, and children and parents who participated our study.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th text revision ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual for mental disorders: DSM-5.* Washington, DC: American Psychiatric Association.

Boucenna, S., Anzalone, S., Tilmont, E., Cohen, D., & Chetouani, M. (2014). Extraction of social signatures through imitation learning between a robot and a human partner. *IEEE Transactions on Autonomous Mental Development*.

Breazeal, C. (2002). *Designing sociable robots.* Cambridge: The MIT Press.

Breazeal, C., Harris, P. L., Desteno, D., Kory Westlund, J. M., Dickens, L., & Jeong, S. (2016). Young children treat robots as informants. *Topics in Cognitive Science, 8*(2), 481–491. https://doi.org/10.1111/tops.12192.

Bushwick, N. L. (2001). Social learning and the etiology of autism. *New Ideas in Psychology, 19*, 49–75. https://doi.org/10.1016/S0732-118X(00)00016-7.

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development, 69*(3), 672–691.

Chandler, M., Fritz, A. S., & Hala, S. (1989). Small scale deceit: Deception as a marker of 2-, 3-, and 4-year-olds’ early theories of mind. *Child Development, 60*(6), 1263–1277. https://doi.org/10.2307/1130919.

Clément, F., Koenig, M., & Harris, P. (2004). The ontogenesis of trust. *Mind & Language, 19*(4), 360–379. https://doi.org/10.1111/j.0268-1064.2004.00263.x.

Constantino, J., & Gruber, C. (2005). *Social responsiveness scale.* Los Angeles: CA: Western Psychological Services.

Corriveau, K., & Harris, P. L. (2009). Choosing your informant: Weighing familiarity and recent accuracy. *Developmental Science, 12*(3), 426–437. https://doi.org/10.1111/j.1467-7687.2008.00792.x.

Costescu, C. A., Vanderborght, B., & David, D. O. (2014). Reversal learning task in children with autism spectrum disorder: A robot-based approach. *Journal of Autism and Developmental Disorders, 45*(11), 3715–3725. https://doi.org/10.1007/s10803-014-2319-z.

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. In Y. Munakata, & M. H. Johnson (Eds.). *Processes of change in brain and cognitive development. Attention and performance XXI* (pp. 249–274). Oxford: Oxford University Press.

Dautenhahn, K., & Billard, A. (2002). *Games children with autism can play with robota, a humanoid robotic doll. Universal Access and assistive Technology*.

DeVries, R. (1970). The development of role-taking as reflected by behavior of bright, average, and retarded children in a social guessing game. *Child Development, 41*(3), 759–770.

Ding, X. P., Heyman, G. D., Fu, G., Zhu, B., & Lee, K. (2018). Young children discover how to deceive in 10 days: A microgenetic study. *Developmental Science, 21*(3), https://doi.org/10.1111/desc.12566.

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems, 42*(3–4), 143–166. https://doi.org/10.1016/S0921-8890(02)00372-X.

Fu, G., Xu, F., Cameron, C. A., Heyman, G., & Lee, K. (2007). Cross-cultural differences in children's choices, categorizations, and evaluations of truths and lies. *Developmental Psychology, 43*(2), 278–293. https://doi.org/10.1037/0012-1649.43.2.278.

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(5), 693–705. https://doi.org/10.1007/s10803-008-0674-3.

Gotham, K., Risi, S., Dawson, G., Tager-flusberg, H., Joseph, R., Carter, A., et al. (2008). A replication of the autism diagnostic observation schedule (ADOS) revised algorithms. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(6), 642–651. https://doi.org/10.1097/CHI.0b013e31816bffb7.

Gratch, G. (1964). Response alternation in children: A developmental study of orientations to uncertainty. *Vita Humana, 7*(1), 49–60. https://doi.org/10.1159/000270053.

Itakura, S., Ishida, H., Kanda, T., Shimada, Y., Ishiguro, H., & Lee, K. (2008). How to build an intentional android: Infants' imitation of a robot's goal-directed actions. *Infancy, 13*(5), 519–532. https://doi.org/10.1080/15250000802329503.

Jaswal, V. K., Croft, A. C., Setia, A. R., & Cole, C. A. (2010). Young children have a specific, highly robust bias to trust testimony. *Psychological Science, 21*(10), 1541–1547. https://doi.org/10.1177/0956797610383438.

Jipson, J. L., & Gelman, S. A. (2007). Robots and rodents: children's inferences about living and nonliving kinds. *Child Development, 78*(6), 1675–1688. https://doi.org/10.1111/j.1467-8624.2007.01095.x.

Kahn, P. H., Friedman, B., Pérezgranados, D. R., & Freier, N. G. (2006). Robotic pets in the lives of preschool children. *Interaction Studies, 7*(3), 1449–1452. https://doi.org/10.1145/985921.986087.

Kahn, P. H., Jr., Gary, H. E., & Shen, S. (2013). Children's social relationships with current and near-future robots. *Child Development Perspectives, 7*(1), 32–37. https://doi.org/10.1111/cdep.12011.

Kahn, P. H., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., Gill, B. T., ... Shen, S. (2012). "Robovie, you'll have to go into the closet now": Children's social and moral relationships with a humanoid robot. *Developmental Psychology, 48*(2), 303–314. https://doi.org/10.1037/a0027033.

Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development, 76*(6), 1261–1277. https://doi.org/10.1111/j.1467-8624.2005.00849.x.

Kozima, H., Nakagawa, C., & Yasuda, Y. (2005). Interactive robots for communication-care: A case-study in autism therapy. *IEEE international workshop on robot and human interactive communication* (pp. 341–346). IEEE. https://doi.org/10.1109/ROMAN.2005.1513802.

Kumazaki, H., Warren, Z., Swanson, A., Yoshikawa, Y., Matsumoto, Y., Ishiguro, H., et al. (2018). Impressions of humanness for android robot may represent an endophenotype for autism spectrum disorders. *Journal of Autism and Developmental Disorders, 48*(2), 632. https://doi.org/10.1007/s10803-017-3365-0.

Lee, K. (2013). Little liars: Development of verbal deception in children. *Child Development Perspectives, 7*, 91–96.

Lee, K., & Cameron, C. A. (2000). Extracting truth information from lies: The emergence of representation-expression distinction in preschool children. *Merrill-Palmer Quarterly, 40*, 1–20.

Leekam, S. R., & Prior, M. (1994). Can autistic children distinguish lies from jokes? A second look at second-order belief attribution. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 35*(5), 901–915. https://doi.org/10.1111/j.1469-7610.1994.tb02301.x.

Lee, J., Takehashi, H., Nagai, C., Obinata, G., & Stefanov, D. (2012). Which robot features can stimulate better responses from children with autism in robot-assisted therapy? *International Journal of Advanced Robotic Systems, 9*(5), 1. https://doi.org/10.5772/51128.

Lewis, M., Stanger, C., & Sullivan, M. W. (1989). Deception in 3-year-olds. *Developmental Psychology, 25*, 439–443. https://doi.org/10.1037/0012-1649.25.3.439.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule—generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*, 205–223. https://doi.org/10.1023/A:1005592401947.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659–685. https://doi.org/10.1007/BF02172145.

Martini, M. C., Gonzalez, C. A., & Wiese, E. (2016). Seeing minds in others - can agents with robotic appearance have human-like preferences? *PLoS One, 11*(1), e0146310. https://doi.org/10.1371/journal.pone.0146310.

Maurer, C., Chambon, V., Bourgeois-Gironde, S., Leboyer, M., & Zalla, T. (2017). The influence of prior reputation and reciprocity on dynamic trust-building in adults with and without autism spectrum disorder. *Cognition, 172*, 1–10. https://doi.org/10.1016/j.cognition.2017.11.007.

Moriguchi, Y., Kanda, T., Ishiguro, H., Shimada, Y., & Itakura, S. (2011). Can young children learn words from a robot? *Interaction Studies, 12*(1), 107–118. https://doi.org/10.1075/is.12.1.04mor.

Palsbo, S. E., & Hood-Szivek, P. (2012). Effect of robotic-assisted three-dimensional repetitive motion to improve hand motor function and control in children with handwriting deficits: A nonrandomized phase 2 device trial. *American Journal of Occupational Therapy Official Publication of the American Occupational Therapy Association, 66*(6), 682. https://doi.org/10.5014/ajot.2012.004556.

Peca, A., Simut, R., Pintea, S., Costescu, C., & Vanderborght, B. (2014). How do typically developing children and children with autism perceive different social robots? *Computers in Human Behavior, 41*(41), 268–277. https://doi.org/10.1016/j.chb.2014.09.035.

Rothbaum, F., Pott, M., Azuma, H., Miyake, K., & Weisz, J. (2000). The development of close relationships in Japan and the United States: Paths of symbiotic harmony and generative tension. *Child Development, 71*, 1121–1142. https://doi.org/10.1111/1467-8624.00214 ·.

Rusca, G., & Tonucci, F. (1992). Development of the concepts of living and animal in the child. *European Journal of Psychology of Education, 7*(2), 151–176. https://doi.org/10.1007/BF03172891.

Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The 'windows task' as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology, 9*(2), 331–349. https://doi.org/10.1111/j.2044-835X.1991.tb00881.x.

Scassellati, B., Admoni, H., & Matarić, M. (2012). Robots for use in autism research. *Annual Review of Biomedical Engineering, 14*(1), 275–294. https://doi.org/10.1146/annurev-bioeng-071811-150036.

Shiomi, M., Kanda, T., Ishiguro, H., & Hagita, N. (2007). Interactive humanoid robots for a science museum. *IEEE Intelligent Systems, 22*(2), 305–312. https://doi.org/10.1109/MIS.2007.37.

Shultz, T. R., & Cloghesy, K. (1981). Development of recursive awareness of intention. *Developmental Psychology, 17*(4), 465–471. https://doi.org/10.1037/0012-1649.17.4.465.

Sodian, B. (1991). The development of deception in young children. *British Journal of Developmental Psychology, 9*(1), 173–188. https://doi.org/10.1111/j.2044-835X.1991.tb00869.x.

Ueyama, Y. (2015). A bayesian model of the uncanny valley effect for explaining the effects of therapeutic robots in autism spectrum disorder. *PLoS One, 10*(9), https://doi.org/10.1371/journal.pone.0138642.

Vanderbilt, K. E., Liu, D., & Heyman, G. D. (2011). The development of distrust. *Child Development, 82*(5), 1372–1380. https://doi.org/10.1111/j.1467-8624.2011.01629.x.

Wechsler, D. (2003). *Manual for the wechsler intelligence scale for children* (4th ed.). San Antonio, TX: The Psychological Corporation.

Xu, F., Bao, X., Fu, G., Talwar, V., & Lee, K. (2010). Lying and truth-telling in children: From conception to action. *Child Development, 81*, 581–596. https://doi.org/10.1111/j.1467-8624.2009.01417.x.

Yang, Y., Tian, Y., Fang, J., Lu, H., Wei, K., & Yi, L. (2017). Trust and deception in children with autism spectrum disorders: A social learning perspective. *Journal of Autism and Developmental Disorders, 47*(3), 1–11. https://doi.org/10.1007/s10803-016-2983-2.

Yi, L., Fan, Y., Li, J., Huang, D., Wang, X., Tan, W., ... Lee, K. (2014). Distrust and retaliatory deception in children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 8*(12), 1741–1755. https://doi.org/10.1016/j.rasd.2014.09.006.