Journal of Autism and Developmental Disorders (2020) 50:1973–1987
https://doi.org/10.1007/s10803-019-03953-x

ORIGINAL PAPER

Outcomes of a Robot-Assisted Social-Emotional Understanding Intervention for Young Children with Autism Spectrum Disorders

Flavia Marino[1] · Paola Chilà[1] · Stefania Trusso Sfrazzetto[1] · Cristina Carrozza[1] · Ilaria Crimi[1] · Chiara Failla[1] · Mario Busà[1] · Giuseppe Bernava[1] · Gennaro Tartarisco[1] · David Vagni[1] · Liliana Ruta[1] · Giovanni Pioggia[1]

Published online: 9 March 2019

Abstract

This study is a randomized control trial aimed at testing the role of a human-assisted social robot as an intervention mediator in a socio-emotional understanding protocol for children with autism spectrum disorders (ASD). Fourteen children (4–8 years old) were randomly assigned to 10 sessions of a cognitive behavioural therapy (CBT) intervention implemented in a group setting either with or without the assistance of a social robot. The CBT protocol was based on Rational Emotive Behaviour Therapy (REBT) principles. Pre- and post-intervention assessments were conducted using the Test of Emotional Comprehension (TEC) and the Emotional Lexicon Test (ELT). Substantial improvements in contextualized emotion recognition, comprehension and emotional perspective-taking through the use of human-assisted social robots were attained.

Keywords Social · Robot · ASD · CBT · Emotion · RCT

Abbreviations

ABC	Antecedent-behaviour-consequence
ADOS-2	Autism diagnostic observation schedule, second edition
ASD	Autism spectrum disorders
CBT	Cognitive behavioural therapy
CG	Control group
CSS	Comparative severity score
DQ	Developmental quotient
EC	Emotion comprehension
ELT	Emotional Lexicon Test
GMDS-ER 2-8	Griffiths Mental Development Scales, Extended Revised: 2 to 8 years
Mdn	Median
RA-CBT	Robot assisted therapy cognitive behavioural therapy
RAT	Robot assisted therapy
RRB	Repetitive and restricted behaviours
REBT	Rational emotional behaviour therapy
RG	Robot group
SA	Social affect
SD	Standard deviation
TEC	Test of emotional comprehension

Liliana Ruta and Giovanni Pioggia contributed equally to the study.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10803-019-03953-x) contains supplementary material, which is available to authorized users.

✉ Giovanni Pioggia
giovanni.pioggia@cnr.it

1 National Research Council of Italy (CNR) – Institute of Applied Sciences and Intelligent Systems "Eduardo Caianiello" (ISASI), via Torre Bianca snc, Istituto Marino di Mortelle, 98168 Messina, Italy

Introduction

Autism spectrum disorders (ASD) are characterized by social communication deficits as well as restricted and repetitive interests and behaviours with perceptual and sensory processing impairments (American Psychiatric Association 2013).

Difficulties in social cognition are present in many children on the autism spectrum[1] regardless of higher cognitive and language abilities. Social cognitive impairment includes deficits in social awareness, perspective-taking and implicit theory of mind (Leekam 2016). Social communication gaps in those children with functional language and without intellectual disability involve lack of integration with typical complementary nonverbal cues and related difficulties in metacommunication, difficulties in understanding social contexts and maintaining reciprocal conversation (Sharma et al. 2018).

Overall, these innate difficulties give rise to impairments in initiating and sustaining typical peer relationships, and thus potentially contribute to the development of negative beliefs about the self, peer rejection and further social withdrawal and isolation (Berenguer et al. 2018; Jones et al. 2018).

Emotional regulation is a key developmental capacity. Effective emotional regulation is associated with social success, academic readiness, and prosocial behaviours, while emotional dysregulation predicts increases in social and behavioural difficulties over time (Berkovits et al. 2017). Children on the autism spectrum have documented emotional regulatory challenges, especially children previously described through the diagnostic label "Asperger syndrome" (De Giambattista et al. 2018) and, more generally, children with mild language and motor delays but without cognitive disability (Rubenstein et al. 2018).

Some of these difficulties are associated with neurological differences and cognitive learning style differences, which are associated with an ASD diagnosis. Other challenges may be associated with ASD-related social learning differences, which impact the nature and effectiveness of social interactions that are geared towards expanding a child's emotional "tool-box".

Cognitive Behaviour Therapy interventions for Autism Spectrum Disorders

Cognitive behaviour therapy (CBT) is a psychotherapeutic approach routinely used in clinical settings to treat a wide range of mental health disorders, including anxiety in children and young people on the autism spectrum (NICE 2013). CBT interventions are short-term, goal-oriented approaches, which aim to support individuals to recognize the interrelation between their thoughts, behaviours and emotions and to develop new ways of thinking about, and coping with, stressful situations (Beck 2011). Increasing evidence supports the use of CBT protocols for individuals on the autism spectrum who have functional language communication abilities.

Rational Emotional Behaviour Therapy (REBT), is a form of CBT aimed to improve multiple skills related to the three dimensions of Emotion Comprehension (EC) described in the developmental model proposed by Pons et al. (2004). The model comprises nine components grouped into three main dimensions, namely, comprehension of the nature of emotions (recognition of basic emotions and appreciation of mixed emotions), the causes of emotion (the role of external causes, desires, beliefs, memory and moral values) and the fact that emotions may be regulated (the awareness that there may be a difference between the emotion experienced and that displayed and that it is possible to regulate the intensity of current emotional experience).

CBT interventions have been adapted to improve core aspects of ASD, such as social-behavioural adaptations and social skills (Hesselmark et al. 2014; Laugeson et al. 2012), as well as co-occurring emotional and behavioural symptoms, such as low mood and low self-esteem, social anxiety, emotional regulation and aggression, etc. (Lopata et al. 2006; Reaven et al. 2012; Santomauro et al. 2016; Scarpa and Reyes 2011; Sofronoff et al. 2005, 2007).

Overall, systematic reviews have demonstrated that CBT interventions show promising evidence of efficacy and are a viable intervention for individuals on the autism spectrum, although the variability of the study designs, the heterogeneity of the participants' clinical presentation and the methodological limitations reduce the generalizability of study findings (Danial and Wood 2013; Weston et al. 2016; Spain et al. 2015, 2017; Lang et al. 2010; Kreslins et al. 2015; Sukhodolsky et al. 2013).

CBT protocols in children and adults on the autism spectrum have been delivered individually, with the involvement of family members or in a group-based format. While individual CBT protocols ensure a better personalization of the therapy objectives and strategies to the specific needs and skills of the single subject (White et al. 2009), group-based CBT interventions take advantage of the influence of social interaction, promoting experience-sharing, improving self-acceptance and supporting insights of both personal strengths and impairments (Hesselmark et al. 2014).

[1] In this paper, we have chosen to use the expression "on the autism spectrum" to describe children with a diagnosis of autism spectrum disorder according to DSM-5 (American Psychiatric Association 2013). This expression is widely accepted by professionals and stakeholders (Kenny et al. 2016) and avoids the person-first (person with autism or ASD) versus identity-first debate (autistic or ASD person). The children involved in our study are too young to choose for themselves how to self-identify and we do not feel comfortable choosing for them. Unless otherwise specified, in the following we will focus on individuals without intellectual disability, with functional language and with a DSM-5 symptom severity level of 1 or 2 on both DSM-5 domains.

Social Robots for Autism Spectrum Disorders

In the last two decades, an emerging field of study has been the integration of classical cognitive and behavioural therapies with new technologies. Recently, Grynszpan et al. (2014) reported a meta-analysis of technology-based intervention studies that used a pre–post design for children on the autism spectrum, including computer programs, virtual reality, and robotics. The overall mean effect size approached the medium magnitude ($d=0.47$) and the results provided support for the continuing development, evaluation, and clinical usage of technology-based intervention. Nevertheless, their findings demonstrated large differences in the level of research design; in particular, all the robotics studies had to be excluded due to either the type of research design or the number of participants.

In spite of its limited clinical validation (Ismail et al. 2018; Scassellati et al. 2012), Robot-assisted autism therapy (RAT) has gained traction in the past few years to become an emerging and promising application area in autism and researchers and clinicians have developed robotic models with a wide range of appearances, features and functional capabilities that draw from expertise in fields such as engineering and clinical psychology (Sartorato et al. 2017).

Robot-mediated intervention studies have demonstrated preliminary evidence of efficacy in different social communication domains such as joint attention, imitation, turn-taking, communicative behaviours (Pennisi et al. 2016; Bird et al. 2007; Kim et al. 2013) and in teaching as an emulated peer (Saadatzi et al. 2018).

A central issue for the still limited clinical validation of RAT is the very small and heterogeneous clinical samples analysed by most of the studies in the field based on exploratory or single-case designs. To our knowledge, only one RCT study (Yun et al. 2017) and one feasibility study (Mengoni et al. 2017) have been conducted. The RCT focused on young children on the autism spectrum and used a robot-mediated, discrete trial teaching protocol to improve basic social skills such as eye contact and facial emotion recognition, while the feasibility study aimed to test the effectiveness of a humanoid robot to support social skills in young children on the autism spectrum.

Nevertheless, no RCTs have been conducted for a robot-assisted intervention embedded into a group-based CBT for young children on the autism spectrum and, more generally, no RCTs with social robots have aimed at the development of higher-order emotional behaviour and cognition, such as emotional perspective-taking.

Objectives and Hypotheses

To begin filling in the gap in RAT for the autism spectrum, we aim to investigate the feasibility and efficacy of using a partially controlled social robot (NAO) in the context of a group-based CBT protocol focused on emotion comprehension and related mentalizing skills, knowledge of a basic emotional lexicon and the ability to attribute appropriate emotions in relation to the context.

Central to our study, the robot exerts a pivotal role as co-therapist by taking turns, directing attention, delivering cues and prompting and reinforcing adequate responses and behaviours in a natural, interactive way.

We hypothesize that children on the autism spectrum randomly assigned to the robot-mediated intervention will show (1) an increase in their emotional competence after the intervention and (2) higher improvements in their socio-emotional understanding skills compared to children on the autism spectrum who will undergo the intervention without the social robot assistance.

Methods

Inclusion Criteria

The inclusion criteria were as follows: (1) between 4 and 8 years of age; (2) clinical diagnosis of ASD based on the DSM-5 criteria from a licensed clinical child neuropsychiatrist; (3) DSM-5 severity scores from mild (level 1) to moderate (level 2) in both social communication and Restricted Interests and Repetitive Behaviours domains; (4) a verbal and performance Developmental Quotient above 70; (5) no current problems with aggressive behaviour or severe oppositional tendency; (6) no hearing, visual, or physical disabilities that would prevent participation in the intervention; (7) not being on psychiatric medication; and (8) not receiving any other intervention directly related to emotional and/or social skills during the trial.

All participants have a previous diagnosis that was further confirmed through the assessment and the consensus of experienced professionals on the research team (i.e., a child neuropsychiatrist and a clinical psychologist).

Participants

Children were recruited as part of an ongoing research programme and tested at our clinical facilities. We enrolled $N=58$ children on the autism spectrum, aged 4 to 8 years, in our research programme. A first screening based on global and verbal DQ (developmental quotient) was implemented and $n=22$ children were eligible and were further assessed; $n=14$ (12 M:2 F) fully met the entry criteria and were enrolled in the present study; $n=3$ children were excluded because they were already receiving interventions directly related to social-communication skills; $n=4$ children were excluded because they were on medications; and $n=1$ child

Fig. 1 Subject recruitment, assignment, and assessment procedures

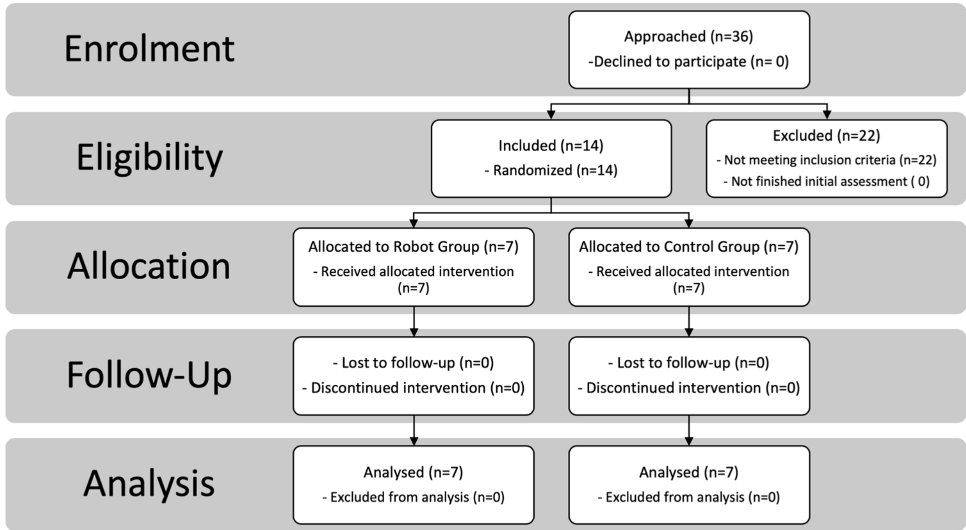

Table 1 Demographic characteristics of the sample

Variable	Unit	Robot group (n = 7)				Control group (n = 7)				Comparison between Groups			
		Mean	SD	Mdn	Full Range	Mean	SD	Mdn	Full Range	*U*	*W*	Z	*p*-value
Age	(months)	73.3	16.1	7.0	48.0–96.0	82.1	12.4	83.0	64.0–96.0	19.0	47.0	−.714	.535
Other therapies	(hours)	2.60	.50	3.00	2.00–3.00	2.60	.50	3.00	2.00–3.00	24.5	52.5	.000	1.00
Formal schooling*	(years)	2.14	.99	2.00	.00–3.00	2.14	.83	2.00	1.00–3.00	25.5	53.5	.136	.902
Weekly schooling*	(hours)	27.0	.0	27.0	27.0–27.0	27.0	.0	27.0	27.0–27.0	24.5	52.5	.000	1.00
Special ed. teacher	(hours)	1.6	.9	1.0	1.0–12.0	1.6	.9	1.0	1.0–12.0	24.5	52.5	.000	1.00
Parents education**	(years)	15.1	2.5	13.0	13.0–18.0	14.9	2.2	13.0	13.0–18.0	26.0	54.0	.218	.902

*The youngest subject (Robot Group) has 48 months, he is in kindergarten for 27 h/week but we counted it as zero in formal schooling variable

**SES was not available for privacy reason, we used parents formal education years as a proxy for SES

was excluded because he did not complete the assessment due to the presence of aggressive behaviours (Fig. 1).

A randomized block design was used to ensure that intervention groups were balanced with respect to gender. $N = 7$ children (6 M:1 F; mean age in months = 73.3; SD = 16.1) were randomly assigned to the intervention group (RG), i.e., a robot-assisted intervention, while $n = 7$ children (6 M:1 F; mean age in months = 82.1; SD = 12.4) were randomly allocated to the control group (CG), applying exactly the same protocol without the use of the assistive-robot.

All included children were attending mainstream public schools for 27 h a week with a special teacher for 10–12 h. Furthermore, they received between 2 and 3 h of intervention outside our facility consisting of a mix of speech and language therapy ($n = 6$ children), occupational therapy ($n = 10$ children) and behavioural intervention ($n = 3$ children). None of the interventions were related to emotional competence or perspective-taking. All parents had a high school or college degree, 15 parents were employees, 5 were freelance professionals and 8 were stay-at-home for their children. All children were Italian (Table 1).

All children scored at or above the clinical cut-off on the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2), module three and scored above the clinical cut-off for developmental delay on the Griffiths Mental Development Scales, Extended Revised: 2 to 8 years (GMDS-ER 2-8). Child psychologists collected information from parents concerning developmental milestones (including joint attention, social interaction, pretend play and repetitive behaviours, with onset prior to 3 years of age) and current behaviours (Table 2). The complete clinical information for each participant is reported in the Supplementary Materials (Table S1).

Ethics

The study was conducted in accordance with the relevant guidelines and regulations for human subjects, and the study design was approved by a National Research Ethics and Bioethics Committee. All the parents of the children who took part in the study signed a written consent form.

Table 2 Clinical characteristics of the sample

Test	Subscale	Robot group (n = 7)				Control group (n = 7)				Comparison between groups			
		Mean	SD	Mdn	Full range	Mean	SD	Mdn	Full range	*U*	*W*	Z	*p*-value
GMDS-ER 2-8	Locomotor	93.6	13.2	98.6	66.0–111	111	19.0	104	89.0-138	13.0	41.0	− 1.47	.165
	Personal-Social	94.6	14.5	93.1	78.9–126	105	14.7	102	89.5–138	12.0	4.0	− 1.60	.128
	Language	94.2	11.5	89.8	79.0–115	87.7	1.9	89.0	68.7–106	29.5	57.5	.640	.535
	Eye and Hand Coordination	94.7	8.4	96.7	79.0–103	104	15.6	98.0	81.4–131	18.0	46.0	− .831	.456
	Performance	90.2	15.7	87.0	66.0–123	110	18.6	97.9	92.4–148	7.0	35.0	− 2.24	.026*
	Practical Reasoning	93.3	14.8	95.8	65.0–117	104	1.5	101	93.0-126	15.0	43.0	− 1.21	.259
	DQ	*95.9*	*13.1*	*91.9*	*82.2–116*	*102*	*1.6*	*95.5*	*92.0-121*	*14.5*	*42.5*	*− 1.28*	*.209*
ADOS-2 module 3	SA	9.29	3.73	8	3–16	8.71	1.48	8	7–12	27.0	55.0	.333	.805
	RRB	5.14	1.73	5	3–9	4.29	2.25	4	1–9	35.5	61.5	1.19	.259
	Total	14.4	4.07	13	8–21	13.0	3.42	12	10–21	32.0	6.0	.972	.383
	CSS	*6.57*	*1.50*	*6*	*4–9*	*6.43*	*1.05*	*6*	*6–9*	*27.5*	*55.5*	*.447*	*.710*

GMDS-ER 2–8 Griffiths Mental Development Scales, Extended Revised: 2 to 8 years, *ADOS-2* Autism diagnostic observation schedule, second edition, *DQ* Developmental Quotient, *SA* Social Affect, *RRB* Restricted Repetitive Behaviours, *CSS* Calibrated Severity Scores, *Mdn* Median, *U* Mann–Whitney U, *W* Wilcoxon W, *SD* Standard Deviation

*p < .05

Experimental Set-up

The sessions were administered in an equipped room called HomeLab. HomeLab is a simulation of a studio flat environment, specifically designed and equipped to enable group-based interaction. The intervention setting included (a) a humanoid robot, NAO, which acted as co-therapist, providing emotional and communication prompts and reinforcements with partial autonomous control, (b) three cameras integrated in a processing module capable of providing real-time behaviour inference data and (c) a widescreen LED TV. The HomeLab setting is pictured in Fig. 2.

The humanoid robot NAO is a commercially available (*SoftBank Robotics*) child-sized, plastic-bodied humanoid robot (58 cm tall, 4.3 kg) equipped with 25 degrees of freedom (DOF), tactile and audio sensors, and a wide variety of actuators (e.g., direct current (DC) servo motors and LED displays).

Group Intervention Protocol

We designed a group-based cognitive behavioural (CBT) emotional understanding skills training protocol based on the principles of REBT. The protocol was specifically designed to be implemented with the assistance of a social robot actively interacting with the children and providing verbal antecedents, prompts and reinforcing consequences according to an Antecedent-Behaviour-Consequence (ABC) model. The protocol was developed to improve multiple skills related to the three dimensions of Emotion Comprehension (EC) described in the developmental model proposed by Pons et al. (2004).

The intervention was delivered in small groups (two groups of 4 children and two groups of 3 children in the RG and CG respectively) and administered by a chartered clinical psychologist and CBT psychotherapist. Clinicians' time was equally allocated between the two groups.

The protocol consisted of 12 sessions (2 for the assessment and 10 for the therapy). Each therapy session lasted 90 min and was administered twice a week, divided into four phases applied through group games and activities.

The first phase consisted of activities focused on emotion recognition skills. The second phase addressed the teaching of context-emotion association. The third phase was centred on discrimination between thoughts and emotions. Lastly, the fourth phase aimed to provide basic strategies for and insights into how to produce and use a repertoire of useful thoughts. The emotion recognition activities were centred

Fig. 2 Collage of the Home Lab setting, NAO robot and an intervention session with the Robot Group

Table 3 Protocol summary

Phase	Week	Focus	Activities	Expected outcome
Initial assessment	1	–	–	–
First	2	emotion recognition	pantomime game	labelling basic emotions
	3			
Second	4	context-emotion association	"how I feel when …?" game	attribute the proper emotion associated with a given situation
	5			
	6			
Third	7	discrimination between thoughts and emotions	"I think-I feel" game	link a thought to an emotion
	8			
	9			
Fourth	10	building a tool-box of useful thoughts	"thought antidote" game	acquisition of useful thoughts related to certain emotional situations
	11			
Final Assessment	12	–	–	–

on a pantomime game, i.e., activities aimed at acquiring the ability to label basic emotions. The emotion and context association tasks are centred on the game of "How I feel when …?", i.e., the presentation of social situations targeted to teach how to attribute the proper emotion associated with the presented situation. The activities related to the emotions to thought recognition were centred on an "I think-I feel" game, i.e., activities aimed at differentiating thoughts from emotions and acquiring the competence to link a thought to an emotion. Lastly, the task dedicated to the learning of a repertoire of useful thoughts was based on activities focused on the acquisition of useful thoughts related to certain emotional situations. In each session, the therapist explained the intended game, asking the participants to repeat the rules in order to memorize them.

A summary of the protocol is reported in Table 3.

Robot Supervision and Autonomous Responses

For this study, we selected NAO for its wide range of vocalization and ease of personalization, sensors and appearance.

A new rule-based supervisory controller was designed within NAO with the capacity to provide emotional and social prompts in the form of body movements and recorded verbal scripts, head and gross-motor orientation of gaze shifts as well as coordinated arm and hand points. Prompts were activated based on real-time data provided back to the robot through a dedicated application running on an iPad driven by the therapist as part of the interaction paradigm.

During each session, the therapist was able to control the robot's behaviours by means of a dedicated application for tablets developed by one of the authors. This application allowed the therapist to choose in advance the list of robot behaviours to be used during the session. In the function of the observed children's responses, the therapist was able to change the robot behaviours in terms of movements and speech in real-time. The therapist was also able to select one of the pre-set response modes for robot reward, encouragement or pause, and perform an action among selected ones in order to stimulate the attention and motivation of the children and to convey positive reinforcements. The robot software was designed to recognize children's faces and eyes, and an automatic routine enabled the robot to maintain eye contact with the children, equally distributing the robot's attention to each child. Moreover, following a specific request from a child, the therapist conveyed "prompts" through the robot, which were driven directly by the therapist. To gradually reduce the conveyed-by-the-robot prompts, a fading technique was implemented, i.e., the prompts were gradually reduced according to the cognitive-behavioural approach. In the RG, such prompts are expressed in the form of "wildcards" proposed by the robot. The child was asked to use the wildcard if he/she had difficulties in providing the answer requested by NAO. Thus, the robot provided the first "prompt". If no one in the group provided the right answer, NAO said "try again", allowing more time to reflect. If a child guessed the right answer within 1 min, NAO conveyed positive feedback (giggles, dancing, applause, etc). If no child provided the correct answer, NAO provided an additional prompt (a "from less to more" prompt hierarchy was adopted). If someone gave the right answer within 1 min, NAO provided positive feedback. If no one guessed the answer, NAO said "This was really difficult. I guess it is…" or alternatively, "I did not even know this…" or "Okay, I would say…" and conveyed the correct answer. For the CG the group, a social skills training programme was conducted using the exact same procedure, with the exception that the role of the robot was replaced by the therapist. In the CG, the therapist took the role of NAO in all the "wildcard" functions and used exactly the same prompts as

Table 4 Pre-intervention outcome measures

Test	Factor	Robot group (n = 7)				Control group (n = 7)				Comparison between groups			
		Mean	SD	Mdn	Full range	Mean	SD	Mdn	Full range	*U*	*W*	Z	*p*-value
TEC	Extrernal	2.29	.700	2	1–3	2.29	.452	2	2–3	25.5	53.5	.145	.902
	Mentalizing	1.43	.495	1	1–2	1.14	.990	1	0–3	3.50	58.5	.830	.406
	Reflexive	.857	.833	1	0–2	1.00	.535	1	0–2	21.5	49.5	−.417	.710
	Total	4.57	1.18	5	3–6	4.14	1.46	4	2–6	28.5	56.5	.526	.620
ELT	*Total*	4.71	.70	5	4–6	4.71	.45	5	4–5	23.5	51.5	−.145	.902

TEC Test of Emotional Comprehension, *ELT* Emotional Lexicon Test, *Mdn* Median, *U* Mann–Whitney U, *W* Wilcoxon W, *SD* standard deviation

well as the fading strategies used by the robot in the RG. No other or subsidiary sources of reinforcement were delivered.

Training and Intervention Fidelity

The therapist who delivered the intervention to both groups was a chartered CBT psychotherapist, well-experienced in working with children on the autism spectrum. She had participated in the protocol development from the beginning and teamed with the engineer to implement the intervention protocol using the robot interface.

Specifically, four one-hour training sessions with the robot engineer were dedicated to the final testing and independent mastery of the tablet-to-robot interface by the therapist.

Adherence to the intervention protocol was assessed using the Cognitive Therapy Scale—Revised (CTS-R) (Blackburn et al. 2001) scoring system (10 scales, each ranging from a minimum of 0 to a maximum of 6). Two external coders, both chartered CBT psychotherapists, rated 50% of the sessions, randomly assigned in both the groups, from video footage.

Therapists' scores on individual items of the CTS-R were all above the score of 4 (Competent/Proficient) in every video scored (total scores, $M = 49.6$, $SD = 4.7$). Inter-rater reliability was adequate, $W = .891$, $p = .005$.

Measures

Outcome Measures

The outcome measures for all of the RG and CG participants were assessed during the weeks before and after the therapy sessions (week 1 and week 12, respectively). The general measures were assessed by direct observations of the children according to the ADOS-2 (Lord et al. 2012) and to the GMDS-ER 2-8 (Luiz et al. 2006). Primary outcome measures were the Test of Emotional Comprehension (TEC) (Pons and Harris 2000) and the Emotional Lexicon Test (ELT) (Grazzani et al. 2009). Both the TEC and the ELT are objective measures of a child's performance related to Emotion Comprehension (EC) and Expression in specific social contexts.

The investigators who assessed the outcome measures were blinded to intervention allocation.

There were no significant differences between groups' pre-intervention outcome measures (Table 4). The complete pre- and post intervention results for each participant are reported in the Supplementary Materials (Tables S2 to S5).

Test of Emotional Comprehension (TEC)

The TEC was standardized by (Pons and Harris 2000) and is considered suitable for testing children on the autism spectrum with a range of ages and abilities, and a useful tool for research (Salomone et al. 2018). According to TEC's theoretical and empirical model (tested with Multidimensional Analysis), it assesses emotional understanding through 9 components that can be grouped into three factors: external (recognition, external causes and belief), mentalist (desire, reminder and hiding) and reflexive (regulation, mixed and morality). For the purpose of this study, we used these three factors as secondary measures.

Emotional understanding is a key aspect of socioemotional competence (Denham et al. 2003, 2010), and includes emotion expression and emotional regulation (Grazzani et al. 2015). ELT was used to evaluate both the basic emotion lexicon repertory and the ability to associate emotions to a specific context.

We administered the standardized Italian version (Albanese and Molina 2008). This test evaluates the comprehension of the nature, causes and regulation of emotion in 3- to 11-year-old children. The test comprises 23 illustrated cards, as some of the nine components under study are assessed using more than one card. The adult reads a short story and shows the child four illustrated faces representing different emotional states—anger, fear, sadness, happiness, or a "neutral" expression – and asks them to select the one that corresponds to the story. For each response assessing one of the nine components, a score of 1 is assigned for a correct answer and a score of 0 for a wrong answer; therefore, the total score ranged from 0 to 9.

Emotional Lexicon Test (ELT)

The ELT was developed to measure the development of emotional lexicon and comprehension in typical children (Gavazzi et al. 2011; Grazzani et al. 2015). Studies in the ASD population have never before been published and this is the first study to our knowledge that uses this measure in children in the autism spectrum. It was included in the research design to assess the pre- and post-intervention differences between the RG and CGs in terms of their linguistic knowledge of the basic emotion lexicon and their ability to match adequate emotions with the context. The measure consists of a series of brief scripted and illustrated scenarios, in each of which an event leads a story character to have an emotional experience. After the story has been read to the child, he or she is invited to choose which of two emotional terms more appropriately defines the protagonist's emotional state and to give a reason for that choice. Given the age of the participants in the present study, only the first part of the test, dedicated to basic emotions, was administered, comprising seven scenarios designed to evaluate the ability to name basic emotions (i.e., fear, joy, sadness, anger, disgust, and shame). The total score awarded to each child thus ranged from 0 to 7.

Statistical Analyses

Considering the small sample size, non-parametric statistics (Mann–Whitney U-tests and the Wilcoxon signed-rank test) were applied in order to analyse the effects of group and intervention.

Inter-rater reliability was calculated using Kendall's Concordance Coefficient (W) for ordinal values.

Total TEC and the ELT scores were used as outcome measures. We used a one-sided test because our hypotheses were directional. We adjusted the alpha level using a Šidák correction for multiple comparison, alpha = .025 for the primary measures, total TEC and ELT scores, alpha = .017 for the three ELT factors, alpha = .006 for the nine ELT scales.

A statistical power analysis was performed for sample size estimation. The analysis was based on clinical judgement due to the absence of similar studies in the literature. To estimate the variables to compute the power, we considered that (1) the different components of the outcome measures are binary (pass/fail); (2) interventions designed to train competences usually aim at a 90% mastery criterion to ensure higher levels of maintenance (Fuller and Fienup 2018); (3) the intervention protocol specifically aims at the different EC components; and (4) usually children on the autism spectrum are similar to the sample (Salomone et al. 2018) lack 1 to 4 (Mdn = 2) of the TEC components targeted by the designed intervention, and our clinical experience suggests that our sample should lack 1 to 2 (Mdn = 2) of the targeted ELT emotions. Therefore, we should expect an average increase of 1.8 points in the total scores (i.e., each participant will gain 1–3 components for each test). From the standardization of the measures, and given the strong homogeneity of the sample, we expected a standard deviation of 1 for the total scores. Therefore, with alpha = .025 and power = 0.8, the projected sample size needed with this effect size (GPower 3.1) is approximately $N = 12$ for this simplest pre- and post-intervention comparison. Thus, our proposed sample size of 14 should be adequate for the main objective of this study.

Our sample was underpowered for parametric tests. Nevertheless, we still wanted to have indicative parametric reference values for future replications. A power analysis for multivariate analysis (MANOVA) and within-subject effect of intervention using repeated measures ANOVA shows that with 7 participants in each group we can identify only very large effects ($f = 0.7$). Therefore, to increase the confidence in the results, we also computed 95% bias-corrected confidence intervals for our data using Bootstrap (n = 1000) with BCa (Bias Corrected Accelerated) and group stratification.

The raw data for each participant comprising demographic and assessment (Table S1), TEC (Table S2) and ELT (Table S3) pre-intervention and post-intervention (Tables S4 and S5 respectively) scores, are provided in the Supplementary Materials in order to give an analytical picture of the sample distribution and to allow replicability.

We reported parametric tests as supplementary analyses. We performed the following: (1) repeated measure ANOVA for intervention gains in both groups using time as a level; (2) MANOVA for between-group comparisons in intervention gains; (3) MANOVA to control for any group effects of the within-group factors on post-intervention behavioural changes, using age, ADOS-2 total, GMDS-ER DQ score, TEC and ELT pre-intervention scores as covariates and post-intervention gains on TEC and ELT as dependent variables (Tables S6 to S8); (4) MANOVA using the different GMDS-ER scales as covariates (Tables S9 and S10); and (5) comparison of pre- and post-intervention passing rates between groups in single TEC (Tables S11 and S12 respectively) and ELT (Tables S13 and S14 respectively) items using a Chi-Squared Test and computing p with 1-sided Fisher's Exact Method.

All of the abovementioned tables are reported in Supplementary Materials.

SPSS software (v. 23, IBM Corporation, Armonk, NY, USA) was used to run statistical analyses.

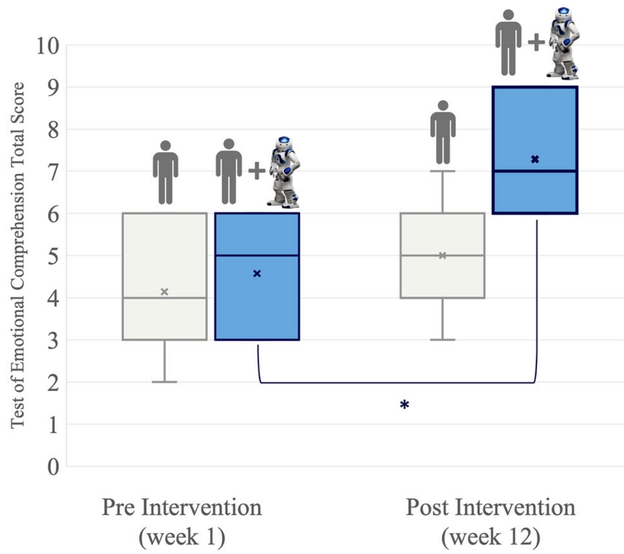

Fig. 3 Box-plot of Post-Intervention Gains for Control and Robot Groups in the Test of Emotional Comprehension. *Footnote*: CG is represented in grey with a man icon) and RG is represented in blue with a man plus a robot. A box is drawn from the first quartile to the third quartile, while the cross is the mean value. The whiskers extend from each quartile to the minimum or maximum. The median line was not depicted because it is over the quartile box (Mdn = 5 for CG at both times, Mdn = 5 for RG pre-intervention and Mdn = 7 for RG)

Results

The demographic and clinical characteristics of the sample are reported in Tables 2 and 3. No significant differences between the RG and the CG were found with respect to any demographic or clinical variable, except for the GMDS-ER performance sub-scale, $U = 7.0$. $W = 35.0$, $Z = -2.24$, $p = .026$, for which the CG displayed significantly higher scores. However, the performance DQ did not predict either TEC, $F(1,6) = .004$, $p = .954$, $b = -.001$, 95% CI [−.171, .148], or ELT $F(1,6) = .002$, $p = .97$. $b = .001$, 95% CI [−2.26, 4.47], via the univariate analysis of the multivariate model with GMDS-ER sub-scales as covariates.

Pre-intervention outcome measures were comparable between the two groups (Table 4). In both TEC and ELT, children in both groups scored significantly lower than expected for their chronological age. For the TEC, 7 children scored 2SD below their age norm, 3 children were 1SD below and 4 children scored within the average range. For ELT, between the ages of 5 and 7 years, the majority of typically developing children reach the test ceiling (7) for basic emotions (88% and 98% respectively), while in our group, the median was 5, full range being 4 to 6.

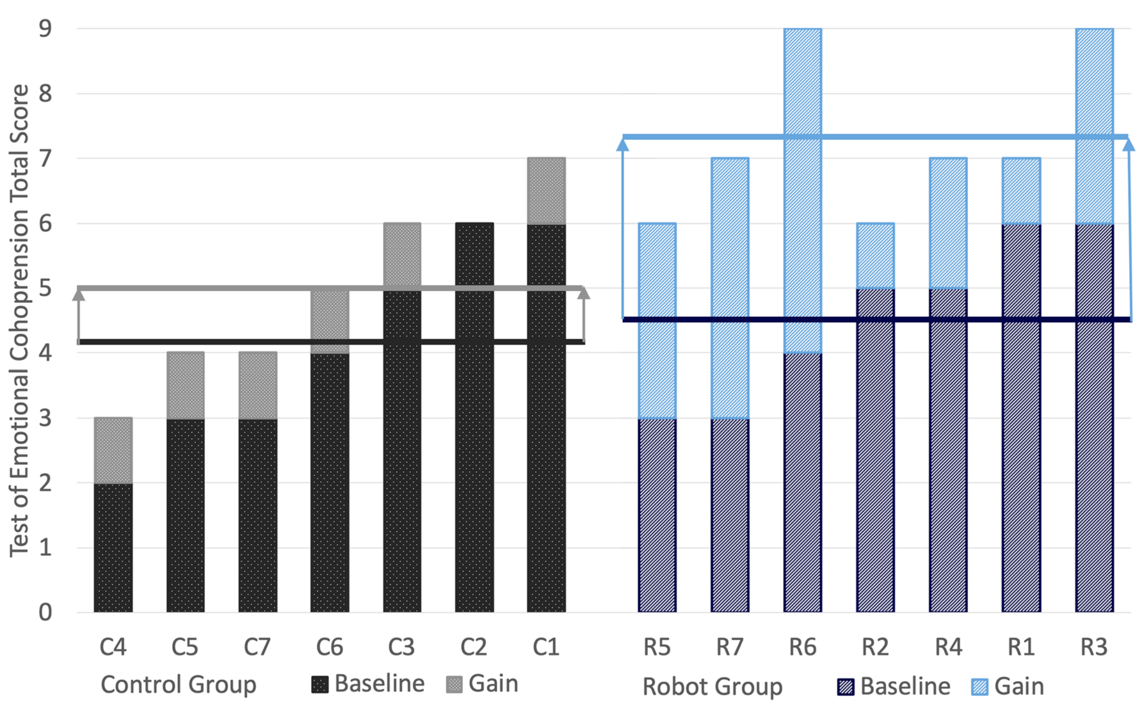

Fig. 4 Single subject gains for the Test of Emotional Comprehension. *Footnote*: Baseline values and gains are in dark and light dotted grey, respectively, for the CG, and dark and light blue stripes, respectively, for the RG. A line was drawn at the averages

Table 5 Pre- and post-intervention comparisons for intervention groups on outcome variables

Group	Test	Factor	Pre intervention (n = 7)				Post intervention (n = 7)				Pre post intervention comparison			
			Mean	SD	Mdn	Full range	Mean	SD	Mdn	Full range	*U*	*W*	Z	*p*-value
RG	TEC	External	2.29	.700	2	1–3	3.00	.000	3	3–3	38.5	66.5	2.26	.037
		Mentalizing	1.43	.495	1	1–2	2.57	.728	3	1–3	42.5	7.5	2.44	.009$^{\#}$
		Reflexive	.857	.833	1	0–2	1.71	.881	1	1–3	36.0	64.0	1.55	.083
		Total	4.57	1.18	5	3–6	7.29	1.16	7	6–9	47.0	75.0	2.93	.001**
	ELT	*Total*	4.71	.70	5	4–6	7.00	.00	7	7–7	49.0	77.0	3.38	.001**
CG	TEC	External	2.29	.452	2	2–3	2.57	.495	3	2–3	31.5	59.5	1.04	.383
		Mentalizing	1.14	.990	1	0–3	1.43	.904	1	0–3	29.0	57.0	.605	.620
		Reflexive	1.00	.535	1	0–2	1.00	.756	1	0–2	24.5	52.5	.000	1.00
		Total	4.14	1.46	4	2–6	5.00	1.31	5	5–6	6.5	32.5	1.04	.318
	ELT	*Total*	4.71	.45	5	4–5	5.29	.452	5	4–5	36.5	64.5	1.93	.128

RG Robot Group, *CG* Control Group. *TEC* Test of Emotional Comprehension, *ELT* Emotional Lexicon Test, *Mdn* Median, *U* Mann–Whitney U, *W* Wilcoxon W, *SD* standard deviation

*p < .025, **p < .005, $^{\#}$p < .017 α-level adjusted using Sidak correction for multiple comparison. $ all the children are at top-score

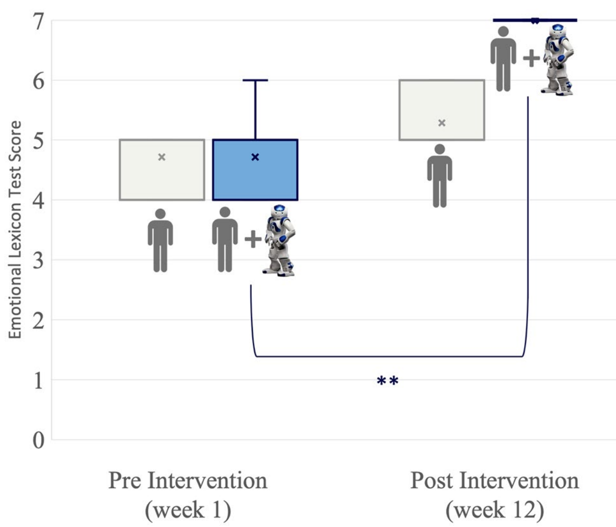

Fig. 5 Box-plot of Post-Intervention Gains for Control and Robot Groups in the Emotional Lexicon Test. *Footnote*: CG is represented in grey with a man icon and RG is represented in blue with a man plus a robot. A box is drawn from the first quartile to the third quartile, while a line is drawn at the median and the cross is the mean value. The whiskers extend from each quartile to the minimum or maximum

Intervention Outcome

All children completed the intervention programme. As shown in Fig. 3, the RG displayed a significant improvement on TEC, $U = 47.0$, $W = 75.0$, $Z = 2.93$, $p = .001$, with a gain of 59% in the total score, and all children in the group improved (Fig. 4). Conversely, the CG, although improving on the TEC total score by 21%, showed no significant change, $U = 32.5$, $W = 60.5$, $Z = 1.04$, $p = .159$ (Table 5).

Likewise, the post-intervention total scores on ELT increased by 48% in the RG and by 12% in the CG (Fig. 5). The Mann–Whitney test indicated a significant improvement in the RG, $U = 49.0$, $W = 77.0$, $Z = 3.38$, $p = .001$, but not in the CG, $U = 36.5$, $W = 64.5$, $Z = 1.92$, $p = .064$ (Table 5). All children in the RG improved (Fig. 6).

Comparing the RG and the CG on intervention gains, a main effect of group was found for both the TEC total score, $U = 43.0$, $W = 71.0$, $Z = 2.62$, $p = .009$, and the ELT total score, $U = 47.0$, $W = 75.0$, $Z = 2.98$, $p = .001$, with the RG showing a significantly higher degree of improvement compared to the CG (Tables 5, 6).

Ancillary Analysis

Via the repeated measures ANOVA, we found a significant main effect of time for TEC, $F(1,12) = 37.5$, $p < .001$, $\eta^2 = .758$, and a significant time x group interaction, $F(1,12) = 10.1$, $p = .008$, $\eta^2 = .458$ with a difference in gain between RG and CG of 1.86 points, 95% CI [.586, 3.12]. We also found a significant main effect of time for ELT, $F(1,12) = 66.7$, $p < .001$, $\eta^2 = .847$, and a significant time x group interaction, $F(1,12) = 24.0$, $p < .001$, $\eta^2 = .667$, with a difference in gain between RG and CG of 1.71 points, 95% CI [.952, 2.48]. For the RG alone there was a significant effect of time both on TEC, $F(1,6) = 23.0$, $p = .003$, $\eta^2 = .793$, with a mean gain of 2.71 points, 95% CI [1.71, 3.86], and on ELT, $F(1,6) = 64.0$, $p < .001$, $\eta^2 = .914$, with a mean gain of 2.29 points, 95% CI [1.86, 2.71]. A significant effect of time was also present in the CG alone on TEC, $F(1,6) = 36.0$, $p = .001$, $\eta^2 = .857$, with a mean gain of .857, 95% CI [.714, 1.00], and on ELT, $F(1,6) = 8.00$. $p = .030$. $\eta^2 = .571$, with a mean gain of .571, 95% [.286, .857].

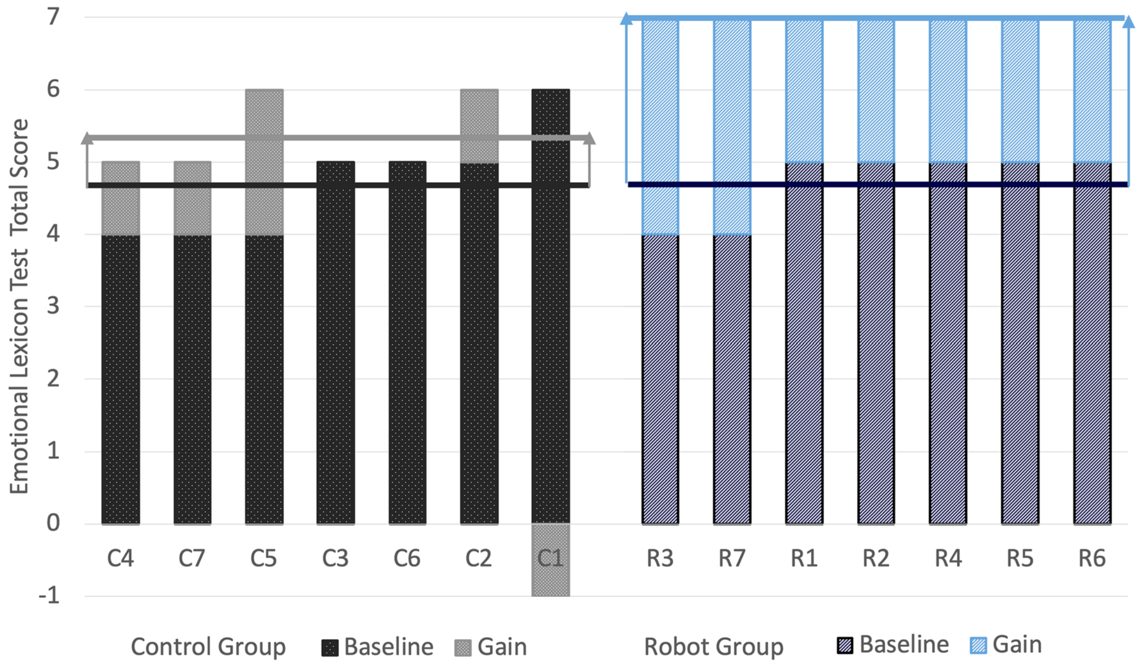

Fig. 6 Single-subject gains for the Emotional Lexicon Test. *Footnote*: Baseline values and gains are in dark and light dotted grey, respectively, for the CG, and dark and light blue stripes, respectively, for the RG. A line was drawn at the averages

Table 6 Comparisons in gains on the outcome measures for the intervention groups

Test	Factor	Robot group (n = 7)				Control group (n = 7)				Comparison between groups			
		Mean	SD	Mdn	Full Range	Mean	SD	Mdn	Full Range	*U*	*W*	Z	*p*-value
TEC	External	.71	.700	1	0–2	.286	.452	0	0–1	32.5	6.5	1.16	.159
	Mentalizing	1.14	.990	1	− 1 to 2	.286	.700	0	− 1 to 1	38.0	66.0	1.82	.049#
	Reflexive	.857	1.12	1	− 1 to 3	.000	.756	0	− 1 to 1	35.5	63.5	1.49	.083
	Total	2.71	1.39	3	1–5	.857	.350	1	0–1	43.0	71.0	2.62	.009*
ELT	*Total*	2.29	.700	2	1–3	.571	.495	1	0–1	47.0	75.0	2.98	.001**

TEC test of emotional comprehension, *ELT* Emotional Lexicon Test, *Mdn* median, *U* Mann–Whitney U, *W* Wilcoxon W, *SD* standard deviation

$^*p < .025$, $^{**}p < .005$, $^{\#}p < .05$ but not significant after α-level adjustment using Sidak correction for multiple comparison

In the analysis of outcome predictors, TEC and ELT gains were used as dependent variables, and age, ADOS-2 total, GMDS-ER DQ, TEC and ELT pre-intervention scores were used as covariates for the MANOVA. In the multivariate model there was only a marginal negative effect of baseline ELT, $F(2,6) = 5.51$, $p = .044$, $\eta_p^2 = 647$. Univariate results show a significant effect only for post-intervention ELT $F(1,7) = 8.88$, $p = .021$, $\eta^2 = .559$. Nevertheless, bootstrap estimation of the 95% CI [− .149, .024] contains a zero.

In the analysis of outcome predictors using GMDS-ER sub-scales as covariates, we found no significant multivariate effects and only a marginally significant univariate effect of the Personal-Social DQ on TEC gain, $F(1,6) = 8.76$, $p = .025$, $\eta^2 = .593$ (Table S10). Furthermore, bootstrap estimation of the 95% CI [− .598, .663] contains a zero.

Analysing the effect of intervention on TEC items, we found no significant change for any specific item (Tables S11 and S12 in the Supplementary Materials). Nevertheless, the *mentalizing* scales show a significant improvement in the RG, $U = 42.5$, $W = 7.50$, $Z = 2.44$, $p = .009$, increasing by 1.14 points, 95% CI [.428, 1.71].

In the RG only, there was a significant and equal gain in *disgust* and *shame* ELT items, X^2 (1, $N = 14$) = 7.78, $p = .010$. The passing rate increased from 28.6 to 100% after the intervention (Table S13).

Discussion

To our knowledge, this is the first study reporting results of a randomized controlled trial of a Robot-Assisted Cognitive Behavioural Therapy (RA-CBT) focused on emotional understanding and social cognition for children on the autism spectrum. The CBT protocol was based on Rational Emotive Behaviour Therapy (REBT) according to the developmental model proposed by Pons et al. (2004) and designed to be implemented through a supervised social robot providing cues, prompts and reinforcements. The intervention consisted of 10 group-based sessions engaging young children in pro-social interactive play activities.

We found a large effect of the social robot in boosting learning of socio-emotional understanding skills, as demonstrated by the significant changes reported by the RG on TEC with a mean gain of 2.71 points (+59.4%), 95% CI [1.71, 3.86], and the ELT with a mean gain of 2.29 points (+48.5%), 95% CI [1.86, 2.71] scores, as well as by the group differences in the magnitude of improvement after intervention, for TEC, 95% CI [.586, 3.12] and ELT, 95% CI [.952, 2.48].

Specifically, an increased ability to understand the beliefs, emotions and thoughts of others was found as confirmed by significant improvement of 1.14 points (+80%), 95% CI [.428, 1.71], in the *Mentalizing* subscale of the TEC in the RG.

Analysing the differences between TEC single items, we found that *Recognition, External Causes* and *Desire* showed no significant changes after intervention. An explanation for this result might be related to a ceiling effect of these outcome measures at baseline. Another explanation might be a possible moderator effect of the trial length and/or the social context should be considered, as reported by Sprung et al. (2015) in their meta-analysis.

Mixed Emotion and *Morality* were not targeted by the intervention and showed no significant improvement trend. *Belief, Regulation* and *Hiding* showed large increases after treatment. However, we did not aim at finding significant differences in the single components; therefore, the study was likely underpowered to find them in those binary variables taken alone. *The Remainder* showed a large increase in the RG alone and led the effect on the *External factor*; however, it was not significant after correction for multiple comparisons. Future replications with larger samples are needed to more precisely underpin the intervention effects on the single TEC subdomains.

For ELT, interestingly, children in the RG showed significant improvements in performance not only in the five basic emotions (anger, disgust, fear, happiness, and sadness) but also for shame (Tables S13 and S14 in the Supplementary Materials). Pre-intervention measures (Table S2) showed that none of the CG children and only 2 of the RG children were able to recognize and explain shame. This is in line with previous literature reporting difficulties for children on the autism spectrum in the recognition and understanding of self-conscious emotions (Heerey et al. 2003). After intervention, all of the children in the RG were able to properly recognize the emotion shame on ELT, while none of the CG children were able to. Interestingly, shame was not explicitly taught during the protocol. We hypothesized the following two possible explanations: (1) robot therapy increased the generalization to other emotions, (2) the presence of two children in the RG who had already passed the item at baseline had led the other children to learn the emotion through incidental social learning. Future studies can shed light on this topic. However, we analysed the data excluding the shame item. One RG participant was excluded from the analysis because he was already at the upper limit. Interestingly, in the RG, all children reached the top score on ELT, learning from one to three new emotions (M = 1.83), while in the CG, no child reached the top score and learned up to one new emotion (M = 0.571). The difference was still significant, $U = 38.0$, $W = 59.0$, $Z = 2.57$, $p = .014$. Furthermore, following the intervention, all children in the RG, but none in the CG, reached the top scores on the ELT.

We had no drop-outs during the intervention. All of the children in the RG demonstrated high interest in the robot and sustained attention and motivation throughout the intervention sessions. The children were exposed to the robot's pro-social behaviours that were usually demonstrated in typical peer-to-peer interactions such as singing with them, talking to them, asking/responding to questions, initiating and sharing play activities. In general, the child-robot interaction was not influenced by technology concerns, e.g., preoccupation with the robot as an object, specific sensory interest, or anxiety. The robot behaviours and the therapist training were optimized to be tuned in with the groups' emotional and social interaction style. As reported by the parents, almost all the children in both the RG and CG spontaneously practised the trained skills at home subsequent to the intervention. This evidence encourages further studies addressing generalization issues and long-term maintenance of robot-assisted interventions.

Limitations

The study has some limitations. First, we should acknowledge that the sample size was small, which may have accounted for the lack of significant improvement shown by the CG in our non-parametric analysis. Nevertheless, the ancillary analysis using repeated measures ANOVA and bias-corrected accelerated bootstrap suggested a small but significant effect also for the CG.

However, we cannot exclude other possible explanations for the smaller effect in the CG. We designed the protocol to be administered with a robot and its specific design could be the cause of the effect. Furthermore, we cannot exclude a more general co-therapist effect. Future studies should compare a therapist plus robot against a therapist plus a human or pet co-therapist, and compare protocols specifically designed for Robot Assisted Therapy and protocols of already proven efficacy using only human therapists. The difference in the complexity of the tasks could be a reason for the difference between CG and RG in our study compared to the similar improvement in both groups reported by the only previous RCT using social robots by Yun et al. (2017). Qualitatively, in our study there was a difference in sustained attention between the conditions. Future studies should quantitatively compare not only the attentional engagement but also the number of exchanges, contingency of reinforcements, consistency of the stimuli and other measures of administration reliability. It is possible that robot assistance could also increase the engagement of a human therapist already involved in the delivery of a complex, and already taxing, intervention, and our reliability measure was not able to capture the difference.

Furthermore, only two girls participated in our study, and although equally assigned to the different intervention conditions, it is not possible to infer any specific possible gender difference in the applicability and effectiveness of social robot mediated intervention. This is a very important and timely research and clinical question to bear in mind, and a specific study designed to address this topic is desirable.

Future Directions

Finally, a comparison of TEC and ELT scores with normative data (Albanese and Molina 2008; Gavazzi et al. 2011) shows that all of the children in RG reached a normative level of emotional comprehension after the intervention, with the three youngest (48–73 months of age) well over the age norm. Salomone et al. (2018) found an age-related increase in emotional understanding skills between children on the autism spectrum and their typical peers, which reflects the challenges of increased social expectation. While increased exposure to social contexts may promote the development of emotion comprehension skills in typical development, if this is not addressed within specific psycho-educational interventions, development is not assured and exposition could be detrimental for children on the autism spectrum. In the absence of intellectual disability, those children can partially compensate for their socio-emotional impairment, but usually with a payback in low self-esteem and internalising symptoms (McCauley et al. 2017). Our study suggests that for emotional comprehension a short robot-assisted cognitive behavioural protocol of only fifteen hours may shift the children's skills to the normative range. Future studies could address the possible protective effect of emotional skills training against the negative social feedback that many children on the autism spectrum may experience during childhood.

Conclusions

The results of the current study supported the idea that a partially controlled assistive robot can be used as an effective augmentative mediator in CBT protocols for socio-emotional understanding in children on the autism spectrum. In line with our hypothesis, we found a significant positive effect of the social robot in fostering children's learning of emotional nature, causes and regulation in relation to specific social contexts. Further studies using larger samples are warranted to replicate these preliminary findings, and all the raw data gathered in this study are reported in the Supplementary Materials for future comparisons.

Acknowledgments We acknowledge all the children and their parents for their help and participation. A special thank you to the Municipality of Messina and our research assistants for their constant support.

Author Contributions FM, LR, DV, PC, STS and GP designed the study. FM and LR performed the diagnostic and functional assessments of the participants. FM, PC and STS administered the protocol. CC, IC and CF administered the TEC and ELT tests as pre- and post-intervention assessments. LR, CC, IC and CF coded the data and reviewed the protocol adherence and fidelity. MB, GB, GT and GP realized the experimental set-up, developed the robot software and trained the therapists. DV, MB, GB, GT and GP processed the experimental data and performed the analysis. FM, LR, DV and GP wrote the manuscript. All authors provided critical feedback and helped shape the research, analysis and manuscript.

Funding This research has been fully supported by the grant from the Sicilian Region of Italy (*Assessorato Regionale dell'Istruzione e della Formazione Professionale, Avviso 11/2017 Rafforzare l'occupabilità nel sistema R&S e la nascita di spin off di ricerca in Sicilia – P.O. FSE 2014/2020*) project no. 2014.IT.05.SFOP.014/3/10.4/9.2.10/0011 - CUP G47B17000100009, entitled SANi.

Data Availability The datasets acquired, used and analyzed during the current study are reported in the Supplementary Materials.

Compliance with Ethical Standards

Ethics Approval The study was conducted in accordance to the relevant guidelines and regulations for human subjects and the study design was approved by the Ethic Committee of the Research Ethics and Bioethics Committee (http://www.cnr.it/ethics) of the National Research Council of Italy (CNR). All the parents of the children who took part in the study gave their assent to participate in this study signing a written consent form.

References

Albanese, O., & Molina, P. (2008). *Lo sviluppo della comprensione delle emozioni e la sua valutazione. La standardizzazione italiana del Test di Comprensione delle Emozioni (TEC)*. Milano: Unicopli.

American Psychiatric Association. (2013). *The diagnostic and statistical manual of mental disorders: DSM-5*. Washington, DC: American Psychiatric Publishing.

Beck, J. (2011). *Cognitive behavioural therapy: Basics and beyond* (2nd edn.). New York: The Guildford Press.

Berenguer, C., Miranda, A., Colomer, C., Baixauli, I., & Roselló, B. (2018). Contribution of theory of mind, executive functioning, and pragmatics to socialization behaviors of children with high-functioning autism. *Journal of Autism and Developmental Disorders, 48*(2), 430–441. https://doi.org/10.1007/s10803-017-3349-0.

Berkovits, L., Eisenhower, A., & Blacher, J. (2017). Emotion regulation in young children with autism spectrum disorders. *Journal of autism and developmental disorders, 47*(1), 68–79.

Bird, G., Leighton, J., Press, C., & Heyes, C. (2007). Intact automatic imitation of human and robot actions in autism spectrum disorders. *Proceedings of the Royal Society B, 274*(1628), 3027–3031. https://doi.org/10.1098/rspb.2007.1019.

Blackburn, I. M., James, I. A., Milne, D. L., Baker, C., Standart, S., Garland, A., & Reichelt F. K. (2001). The revised cognitive therapy scale (CTS-R): Psychometric properties. *Behavioural and Cognitive Psychotherapy, 29*(4), 431–446.

Danial, J. T., & Wood, J. J. (2013). Cognitive behavioral therapy for children with autism. *Journal of Developmental & Behavioral Pediatrics, 34*(9), 702–715. https://doi.org/10.1097/DBP.0b013e31829f676c.

De Giambattista, C., Ventura, P., Trerotoli, P., Margari, M., Palumbi, R., & Margari, L. (2018). Subtyping the autism spectrum disorder: Comparison of children with high functioning autism and asperger syndrome. *Journal of autism and developmental disorders, 41*(1), 138–150.

Denham, S. A., Bassett, H. H., & Wyatt, T. (2010). The socialization of emotional competence. *New Directions for child and adolescent development, 128 (2010)*, 29–49.

Denham, S. A., Blair, K. A., DeMulder, E., Levitas, J., Sawyer, K., Auerbach-Major, S., & Queenan, P. (2003). Preschool emotional competence: Pathway to social competence? *Child Development, 74*(1), 238–256. https://doi.org/10.1111/1467-8624.00533.

Fuller, J. L., & Fienup, D. M. (2018). A preliminary analysis of mastery criterion level: Effects on response maintenance. *Behavior Analysis in Practice, 11*(1), 1–8. https://doi.org/10.1007/s40617-017-0201-0.

Gavazzi, I. G., Ornaghi, V., & Piralli, F. (2011). Teoria della mente e comprensione del lessico psicologico nei bambini: Dati preliminari di validazione del test di lessico emotivo (TLE). *Psicologia Clinica Dello Sviluppo, 15*(1), 257–266.

Grazzani, I., Ornaghi, V., & Piralli, F. (2009). TEST DI LESSICO EMOTIVO.

Grazzani, I., Ornaghi, V., & Riva Crugnola, C. (2015). Emotion comprehension and attachment: A conversational intervention with school-aged children. *Revue Europeenne de Psychologie Appliquee, 65*(6), 267–274. https://doi.org/10.1016/j.erap.2015.10.004.

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

Heerey, E. A., Keltner, D., & Capps, L. M. (2003). Making sense of self-conscious emotion: Linking theory of mind and emotion in children with autism. *Emotion, 3*(4), 394–400. https://doi.org/10.1037/1528-3542.3.4.394.

Hesselmark, E., Plenty, S., & Bejerot, S. (2014). Group cognitive behavioural therapy and group recreational activity for adults with autism spectrum disorders: A preliminary randomized controlled trial. *Autism*. https://doi.org/10.1177/1362361313493681.

Ismail, L. I., Verhoeven, T., Dambre, J., & Wyffels, F. (2018). Leveraging robotics research for children with autism: A review. *International Journal of Social Robotics*, 2018:1–22.

Jones, C. R. G., Simonoff, E., Baird, G., Pickles, A., Marsden, A. J. S., Tregay, J., Charman, T. (2018). the association between theory of mind, executive function, and the symptoms of autism spectrum disorder. *Autism Disorder, 11*(1), 95–109. https://doi.org/10.1002/aur.1873.

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism, 20*(4), 442–462.

Kim, E. S., Berkovits, L. D., Bernier, E. P., Leyzberg, D., Shic, F., Paul, R., & Scassellati, B. (2013). Social robots as embedded reinforcers of social behavior in children with autism. *Journal of Autism and Developmental Disorders, 43*(5), 1038–1049. https://doi.org/10.1007/s10803-012-1645-2.

Kreslins, A., Robertson, A. E., & Melville, C. (2015). The effectiveness of psychosocial interventions for anxiety in children and adolescents with autism spectrum disorder: A systematic review and meta-analysis. *Child and Adolescent Psychiatry and Mental Health, 9*(1), 22. https://doi.org/10.1186/s13034-015-0054-7.

Lang, R., Regester, A., Lauderdale, S., Ashbaugh, K., & Haring, A. (2010). Treatment of anxiety in autism spectrum disorders using cognitive behaviour therapy: A systematic review. *Developmental Neurorehabilitation, 13*(1), 53–63.

Laugeson, E. A., Frankel, F., Gantman, A., Dillon, A. R., & Mogil, C. (2012). Evidence-based social skills training for adolescents with autism spectrum disorders: The UCLA PEERS program. *Journal of Autism and Developmental Disorders, 42*(6), 1025–1036.

Leekam, S. (2016). Social cognitive impairment and autism: what are we trying to explain? *Philosophical Transactions of The Royal Society B*, https://doi.org/10.1098/rstb.2015.0082.

Lopata, C., Thomeer, M. L., Volker, M. A., & Nida, R. E. (2006). Effectiveness of a cognitive-behavioral treatment on the social behaviors of children with Asperger disorder. *Focus on Autism and Other Developmental Disabilities, 21*(4), 237–244.

Lord, C., Rutter, M., Dilavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism diagnostic observation schedule (ADOS-2): Manual* (2nd edn.). Los Angeles: Western Psychological Services.

Luiz, D., Faragher, B., Barnard, A., Knoesen, N., Kotras, N., Burns, L. E., & Challis, D. (2006). *Griffiths mental development scales, extended revised: 2 to 8 years*. Oxford: Hogrefe.

McCauley, J. B., Harris, M. A., Zajic, M. C., Swain-Lerro, L. E., Oswald, T., McIntyre, N., & Solomon, M. (2017). Self-esteem, internalizing symptoms, and theory of mind in youth with autism spectrum disorder. *Journal of Clinical Child & Adolescent Psychology*. https://doi.org/10.1080/15374416.2017.1381912.

Mengoni, S. E., Irvine, K., Thakur, D., Barton, G., Dautenhahn, K., Guldberg, K., & Sharma, S. (2017). Feasibility study of a randomised controlled trial to investigate the effectiveness of using a humanoid robot to improve the social skills of children with autism spectrum disorder (Kaspar RCT): A study protocol. *British Medical Journal Open, 7*(6), 1–9. https://doi.org/10.1136/bmjopen-2017-017376.

NICE. National Institute for Health and Clinical Excellence. (2013). Autism spectrum disorder in under 19s: Support and management. https://www.nice.org.uk/guidance/cg170.

Pennisi, P., Tonacci, A., Tartarisco, G., Billeci, L., Ruta, L., Gangemi, S., & Pioggia, G. (2016). Autism and social robotics: A systematic review. *Autism Research, 9*(2), 165–183. https://doi.org/10.1002/aur.1527.

Pons, F., & Harris, P. (2000). *Test of emotion comprehension: TEC*. Oxford: University of Oxford.

Pons, F., Harris, P. L., & de Rosnay, M. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology, 1*(2), 127–152. https://doi.org/10.1080/17405620344000022.

Reaven, J., Blakeley-Smith, A., Culhane-Shelburne, K., & Hepburn, S. (2012). Group cognitive behavior therapy for children with high-functioning autism spectrum disorders and anxiety: A randomized trial. *Journal of Child Psychology and Psychiatry, 53*(4), 410–419.

Rubenstein, E., Wiggins, L. D., Schieve, L. A., Bradley, C., DiGuiseppi, C., Moody, E.,. & Pence, B. W. (2018). Associations between parental broader autism phenotype and child autism spectrum disorder phenotype in the Study to Explore Early Development. *Autism*. https://doi.org/10.1177/1362361317753563.

Saadatzi, M. N., Pennington, R. C., Welch, K. C., & Graham, J. H. (2018). Small-group technology-assisted instruction: Virtual teacher and robot peer for individuals with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 48*(11), 3816–3830.

Salomone, E., Bulgarelli, D., Thommen, E., Rossini, E., & Molina, P. (2018). Role of age and IQ in emotion understanding in autism spectrum disorder: Implications for educational interventions. *European Journal of Special Needs Education, 6257*, 1–10. https://doi.org/10.1080/08856257.2018.1451292.

Santomauro, D., Sheffield, J., & Sofronoff, K. (2016). Depression in adolescents with ASD: A pilot RCT of a group intervention. *Journal of Autism and Developmental Disorders, 46*(2), 572–588.

Sartorato, F., Przybylowski, L., & Sarko, D. K. (2017). Improving therapeutic outcomes in autism spectrum disorders: Enhancing social communication and sensory processing through the use of interactive robots. *Journal of Psychiatric Research, 9*, 1–11. https://doi.org/10.1016/j.jpsychires.2017.02.004.

Scarpa, A., & Reyes, N. M. (2011). Improving emotion regulation with CBT in young children with high functioning autism spectrum disorders: A pilot study. *Behavioural and Cognitive Psychotherapy, 39*, 495–500. https://doi.org/10.1017/S1352465811000063.

Scassellati, B., Henny Admoni, & 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.

Sharma, S. R., Gonda, X., & Tarazi, F. I. (2018). Autism spectrum disorder: Classification, diagnosis and therapy. *Pharmacology and Therapeutics*. https://doi.org/10.1016/j.pharmthera.2018.05.007.

Sofronoff, K., Attwood, T., & Hinton, S. (2005). A randomised controlled trial of a CBT intervention for anxiety in children with Asperger syndrome. *Journal of Child Psychology and Psychiatry, 46*(11), 1152–1160.

Sofronoff, K., Attwood, T., Hinton, S., & Levin, I. (2007). A randomized controlled trial of a cognitive behavioural intervention for anger management in children diagnosed with Asperger syndrome. *Journal of Autism and Developmental Disorders, 37*(7), 1203–1214.

Spain, D., Sin, J., Chalder, T., Murphy, D., & Happe, F. (2015). Cognitive behaviour therapy for adults with autism spectrum disorders and psychiatric co-morbidity: A review. *Research in Autism Spectrum Disorders, 9*, 151–162.

Spain, D., Sin, J., Harwood, L., Mendez, M. A., & Happé, F. (2017). Cognitive behaviour therapy for social anxiety in autism spectrum disorder: A systematic review. *Advances in Autism, 3*(1), 34–46.

Sprung, M., Münch, H. M., Harris, P. L., Ebesutani, C., & Hofmann, S. G. (2015). Children's emotion understanding: A meta-analysis of training studies. *Developmental Review, 37*, 41–65. https://doi.org/10.1016/j.dr.2015.05.001.

Sukhodolsky, D. G., Bloch, M. H., Panza, K. E., & Reichow, B. (2013). Cognitive-behavioral therapy for anxiety in children with high-functioning autism: A meta-analysis. *Pediatrics, 132*(5), e1341–e1350. https://doi.org/10.1542/peds.2013-1193.

Weston, L., Hodgekins, J., & Langdon, P. E. (2016). Effectiveness of cognitive behavioural therapy with people who have autistic spectrum disorders: A systematic review and meta-analysis. *Clinical Psychology Review, 49*, 41–54. https://doi.org/10.1016/j.cpr.2016.08.001.

White, S., Oswald, D., Ollendick, T., & Scahill, L. (2009). Anxiety in children and adolescents with autism spectrum disorders. *Clinical Psychology Review, 29*(3), 216–229. https://doi.org/10.1016/j.cpr.2009.01.003.Anxiety.

Yun, S.-S., Choi, J., Park, S.-K., Bong, G.-Y., & Yoo, H. (2017). Social skills training for children with autism spectrum disorder using a robotic behavioral intervention system. *Autism Research, 10*(7), 1306–1323. https://doi.org/10.1002/aur.1778.