Gdańsk University of Technology
Faculty of Electronics, Telecommunications and Informatics

TECHNICAL REPORT

1/2021

SYSTEMATIC LITERATURE REVIEW - ROBOT-BASED INTERVENTION FOR CHILDREN WITH AUTISM SPECTRUM DISORDER

Katrin D. Bartl-Pokorny[1,2]*, Małgorzata Pykała[3]*, Duygun Erol Barkana[4], Alice Baird[1], Hatice Köse[5], Tatjana Zorcec[6], Ben Robins[7], Björn W. Schuller[1,8], Agnieszka Landowska[3#]

[1]EIHW – Chair of Embedded Intelligence for Health Care and Wellbeing, University of Augsburg, Augsburg, Germany
[2]Research Unit iDN – interdisciplinary Developmental Neuroscience, Division of Phoniatrics, Medical University of Graz, Austria
[3]Department of Software Engineering, Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, Gdansk, Poland
[4]Yeditepe University, Electrical and Electronic Engineering, Istanbul, Turkey
[5]Faculty of Computer and Informatics, Istanbul Technical University, Istanbul, Turkey
[6]University Children's Hospital, Skopje, North Macedonia
[7]School of Computer Science, University of Hertfordshire, Hatfield, UK
[8]GLAM – Group on Language, Audio, & Music, Imperial College London, London, United Kingdom
*These authors contributed equally to this work.
#Correspondence to: Agnieszka Landowska, PhD, DSc, Eng, GUT professor: nailie@eti.pg.edu.pl

This publication was supported in part by the Erasmus Plus project of European Commission: EMBOA, Affective loop in Socially Assistive Robotics as an intervention tool for children with autism, contract no 2019-1-PL01- KA203-065096. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein. This publication is distributed free of charge.

1. Literature review process

We performed a systematic literature review (SLR) on robot-based interventions targeting emotion-related skills for children with autism spectrum disorder (ASD). According to guidelines proposed by Kitchenham et al. [1], we started with defining research questions. We focused on psychological as well as technical aspects, related to the use of robots in therapy of children with ASD. We finished up with seven research questions, and a summary of the SLR scheme is provided in Table 1, followed by a more detailed explanation.

Table 1. Summary of SLR step-by-step. Abbreviations: PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RQ = Research question; yellow color: Papers with total rank between 4 and 7 were classified to the next phase of analysis; green color: Papers with a total rank equal to or greater than 4 and those which in the previous stage obtained a total rank of 8 were qualified for full text analysis.

Step	Results
1. Setting up research questions	RQ1: What robots are used in interventions in children in autism? RQ2: What emotions are of interest in children with autism? RQ3: What emotion-related skills are addressed in interventions using robots? RQ4: What are challenges of emotional natura in child-robot interaction in autism? RQ5: Are there any recommendations/lessons learned regarding interventions in autism using robots?

2. Defining keywords and search string	`(emotion OR affective OR emotional OR affect)`
	`AND (autism OR ASD OR ASC OR autistic OR "pervasive disorder")`
	`AND (children OR child OR young)`
	`AND (robot OR SAR OR RAT OR robotic OR humanoid)`
	`AND (intervention OR learn OR learning OR teach OR teaching OR tutoring OR therapy OR coaching)`

3. Decisions on search engines, inclusion and exclusion criteria	Scopus, Web of Science, PubMed Inclusion criteria: all years, all type of document, papers that focus on using robots in intervention for children with autism spectrum disorder, papers that report trials with at least 1 child, emotion-related skills are addressed in interventions Exclusion criteria: full text of the paper is not available, the paper is written in a language other than English
4. Data extraction	All three databases, by topic, resulted in 440 papers (312 after removal of duplicates)
5. Multiple-phase selection based on quality criteria and research question	Manual annotating by 5 independent annotators - 4 per tagging stage. Scale: 2 (relevant), 1(not sure), 0(not relevant) Papers with total score 8: 31 Papers with total score 7: 8 Papers with total score 6: 10

	Papers with total score 5: 27 Papers with total score 4: 60 Papers with total score 3: 21 Papers with total score 2: 17 Papers with total score 1: 41 Papers with total score 0: 97 Qualified to reading abstract after title tagging: 105. Papers with total score 8: 15 Papers with total score 7: 3 Papers with total score 6: 7 Papers with total score 5: 3 Papers with total score 4: 23 Papers with total score 3: 18 Papers with total score 2: 16 Papers with total score 1: 5 Papers with total score 0: 15 Qualified to reading full text after title and abstract tagging: 82.
6. Final selection of papers and snowballing technique	15 additional papers came from snowballing technique (7 after removing duplicates) 82+7=89 papers to read 53 qualified for detailed analysis and reporting
7. Extraction of the key findings	For each paper we tag: - number of children in the study (with ASD or typically developing), - wording about children (using expression "child with autism" vs "autistic child"), - emotions analysed or discovered, encountered, by surprise or intentionally, - emotion-based skills trained in interventions with the use of a robot, - robots used. We also note challenges/recommendations or other relevant observations.
8. Reporting systematic review	Apart from this technical report a publication following the PRISMA statement is planned [2].

For our previous SLR on emotion recognition in children with ASD [3], we used the following seven scientific databases: ACM Digital Library, Elsevier Science direct, IEEE Xplore, Scopus, SpringerLink, Web of Science, and PubMed. About 80 percent of search results came from three of them: Scopus, Web of Science, and PubMed. Therefore, we chose these three databases for the present literature search. Scopus and Web of Science are among the most popular scientific databases in technical sciences. PubMed is among the most important for scientists in medical and psychological domains.

In order to obtain optimal search results, we considered fields in databases such as title and topic. Number of obtained results for each database and search field is presented in Table 2. For PubMed we present two numbers for the topic field. The one in parenthesis is number of records while considering all fields. We decided to use it because it is a number comparable with the values for the topic field for other databases and the other one is equal to 0.

Table 2. Number of obtained results for the specified query

#		Search field	
		Title	Topic
Database	Scopus	13	183
	Web of Science	11	176
	PubMed	1	0 (81)

Finally, we decided to use the topic field only, because 25 records is definitely not enough.

After collecting results from different databases, we unified the file format and merged all records to remove duplicates and perform initial analysis. First, papers were evaluated for their title and relevance to research questions by four annotators. The relevance was measured on the three-points scale: 0 – irrelevant, 1 – somehow relevant, 2 – strongly relevant. Papers with total rank between 4 and 7 were classified to the next phase of analysis (marked in yellow in Table 1). These papers were reassessed for relevance to the research questions based on the content of their abstracts. The number of annotators and the relevance scale remained the same. Documents with a total rank equal to or greater than 4 and those which in the previous stage obtained a total rank of 8 were qualified for full text analysis (marked in green in Table 1). Papers scored less than 4 were excluded automatically.

We qualified 82 papers for reading full text after title and abstract tagging. We added 15 additional papers by a snowballing technique. We derived most of them from the previous SLR on emotion recognition in children with ASD [3]. These are articles that did not meet all inclusion criteria in the previous SLR, but may include answers to the research questions of this SLR. We removed 8 of the 15 additional articles as they were already among the 82 papers qualified after the selection, resulting in a final number of 89 publications that qualified for reading full text.

From 89 papers only 53 were qualified for detailed analysis and reporting. We excluded 6 papers - 4 were unavailable, 2 were written in French. Other 30 papers were excluded after reading because there were no children with ASD, no robot was used, no emotion-related skills are addressed in interventions and no emotions monitored in human-robot interaction. However, we agreed to also include papers that do not satisfy these criteria, if they bring some value like a well-described challenge or guideline. Results from this phase of SLR are presented in Figure 1.

We prepared forms in spreadsheets for the extraction of the key findings in interventions regarding robots and emotional skills in children with ASD. We were interested in many aspects, e.g., which emotion-related skills are addressed in interventions using robots, which robots are used, if and which emotions are monitored in human-robot interaction. We made notes about the number of children in the study (with ASD or typically developing), their gender and age. We also wanted to note all challenges, recommendations and other relevant observations reported in the articles. We agreed that each annotator could extend the forms, i.e. by adding new emotions or skills that were not considered at the beginning of this study.

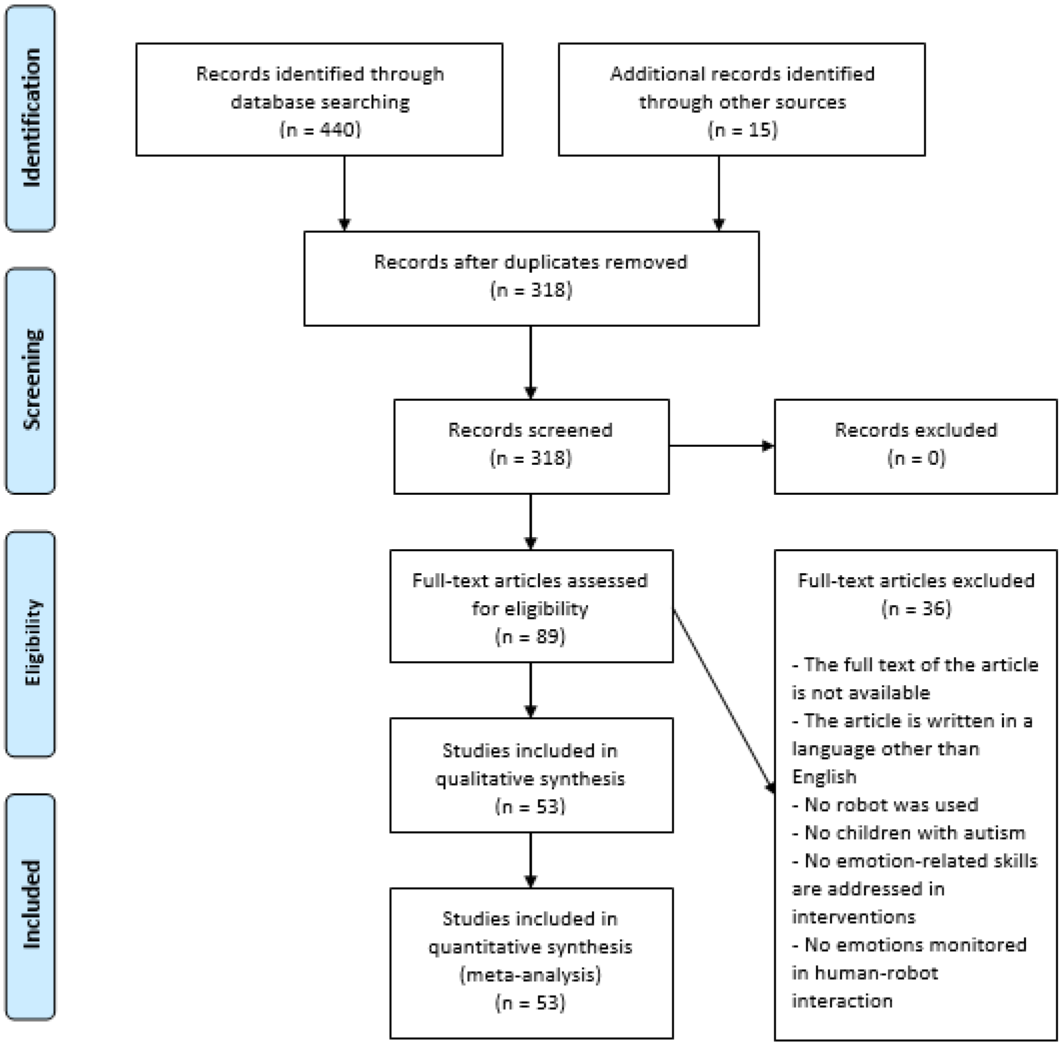

Figure 1. PRISMA flow diagram for the systematic review.

2. Study participants

2.1. Children with ASD vs. typically developing children

Forty-one of 53 articles included participants. One study (T25) included adults only. The other 40 studies investigated children. Thirty-nine of these 40 studies included children with ASD whereas eleven of these studies included typically developing controls. One study (T47) included one typically developing child, but no children with ASD. The total number of participants included by the 41 studies with participants is 756 (min: 1, max: 137, mean: 18.44, median: 10, standard deviation: 25; rounded to two decimals). Of these, 445 were children with ASD and 311 were typically developing participants (137 of these were adults; T25). Twenty studies included 9 or fewer participants respectively; twenty-one studies included 10 or more participants respectively.

2.1.1. Subgroups of participants with ASD

Twenty-five studies included in this SLR investigated participants with ASD and did not specify in their articles whether these children belonged to certain subgroups of ASD. Some studies, however, described their participants by reporting them to belong to ASD subgroups (Figure 2). The most frequently included subgroup is the 'high functioning' subgroup, followed by the 'low functioning' subgroup. It remains open whether the subgroups 'Asperger' and 'high functioning' and the subgroups 'ASD + Intellectual Disability' and 'low functioning' have the same inclusion criteria respectively and could thus be grouped together.

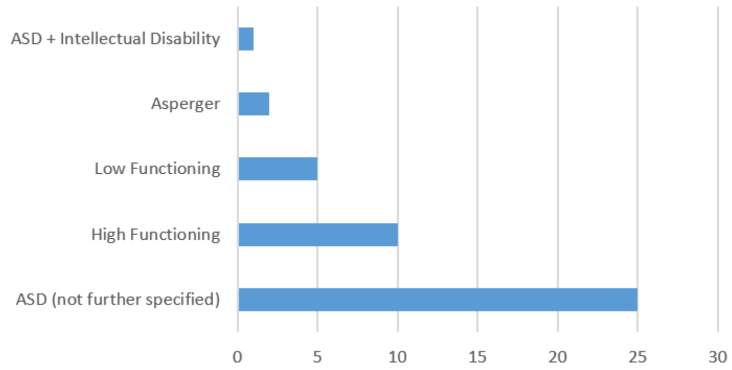

Figure 2. Number of articles including participants of ASD subgroups and participants with ASD without assignment to subgroups.

2.2. Gender

Thirty of the 39 studies including participants with ASD provided information about the gender of the participants with ASD: The gender of a total of 390 participants with ASD was reported. Of these, ten studies included only male participants. From all studies together, 323 participants with ASD were reported to be males (min: 1, max: 52, mean: 10.77, median: 8.5, standard deviation: 11.23) and 67 participants with ASD were reported to be females (min: 0, max: 9, mean: 2.23, median: 2, standard deviation: 2.49). Of the 20 studies including both male and female participants, two studies included an equal amount of female and male participants (T19, T55), two studies included more females than males (T68 , S7), and the remaining 16 studies included more males than females.

2.3. Age

The adult participants of T25 were between 21 and 63 years old (mean: 34.5, standard deviation: 9.6). Thirty-five of the 40 studies including children provided information on the age of the participants. Of these, 30 studies provided the exact ages or the age ranges of the participants. The participants of these 30 studies were between 2 and 20 years old: The youngest participant of the studies was 2-14 years (rounded to full years) respectively (mean: 5.83, median: 5, standard deviation: 3.02); the oldest participant of the studies was 2-20 years (rounded to full years) respectively (mean: 10.47, median: 11, standard deviation: 4.32). Two of the 30 studies included only one participant; for the remaining 28 studies we calculated the number of years between the youngest and the oldest child: it was 0 to 13 years respectively (mean: 4.96, median: 4, standard deviation: 3.4). The five studies that did not provide age ranges either provided the mean age of the participants (T2: 9.03 years, T17: 10 years, T27: 11.4 years, T77: 2.5 years), or the mean age and the standard deviation (S4: 5.4±1.5 years).

2.4. Wordings used to refer to participants

Articles used different wordings to refer to participants with ASD and TD participants (Table 3). The most often used wording for children with ASD was 'child(ren) with ASD' and the often used wording for TD children was 'typically developing child(ren)'.

The wording 'child(ren) with ASD' was used by articles published from the years 2008 to 2020 (median: 2017). The wording 'child(ren) with autism' was used by articles published between 2002 and 2019 (median: 2017). 'Autistic child(ren)' was used by articles published from the years 2002 to 2020 (median: 2015). 'Child(ren) on/with autism spectrum' was used in articles published in 2017 and 2019. The wording 'autism child(ren)' was used by an article published in 2018. The wording 'typically developing child(ren)' was used by articles published between 2014 and 2019 (median: 2018). 'Neurotypical child(ren)' was used in articles published between 2015 and 2019. 'Normal child(ren)' was used by articles published between the years 2012 and 2018. The wording 'typical individual(s)' was used by an article published in 2008.

Table 3. Articles using the given wordings to refer to the participants. Some articles are listed twice or more often because they used two or more wordings.

Wording	Articles
Child(ren) with ASD	T2, T5, T6, T8, T9, T13, T14, T16, T17, T20, T21, T27, T28, T29, T32, T36, T37, T40, T63, T72, S3, S4, S7
Child(ren) with autism	T8, T19, T23, T25, T26, T47, T63, S2, S4, S7, S8
Autistic child(ren)	T2, T18, T41, T42, T47, T73, S2, S3, S4
Child(ren) on/with autism spectrum	T9, T24
Autism child(ren)	T73
Typically developing child(ren)	T5, T6, T14, T20, S4, S7, S8
Neurotypical child(ren)	T8, T13, T17, T42
Normal child(ren)	T41, T47, T73, S3
Typical individual(s)	T63

3. Robots

3.1. Types of robots and their popularity

The articles referred to 32 robots. We have divided them into five categories: humanoid, animal/creature, mobile robot, ball-shaped robot, and other. There are 14 robots in the humanoid category. Various versions of the Darwin-Mini robot from Robotis were used in the research: the Darwin OP (T34) and its newer version, the Darwin OP2. Another humanoid robot listed in the table has appeared in articles under several names: Zeno R-50 (T42), ZECA

(T49, S7), Robokind Zeno R25 (T40). We decided to call it "ZECA" as it stands for "Zeno Engaging Children with Autism".

Five robots, physically resembling animals, and three non-animals, with characteristics indicating that they were intended to resemble a living creature, were assigned to the "animal / creature" category. The animal-like robots included the dinosaur (dinosaur robot Pleo), penguin (PABI), monkey (SAM), parrot (Semi-autonomous KiliRo robot) and dog (Zoomer dog). The other robots assigned to this category are: a small, yellow snowman (Keepon), a creature resembling a droplet with an eye (Muu), and a creature resembling an elephant or Alf from the sitcom of the same title (Probo).

The next category includes six mobile robots. Five of them are remote-controlled robots. Three of these five, in addition to the possibility of remote control, have additional functions such as blowing bubbles to attract child's attention and maintain engagement (Bubble blower), displaying facial expressions of emotions on an attached mobile device (ROMO), moving its arms and emitting noise to check the expression of negative emotions in facial expression and body posture in response to fear (Remote-controlled robot). The sixth robot (Labo-1) is equipped with infrared sensors, as well as a heat sensor. Using these sensors, the robot is able to avoid obstacles and can follow a heat source such as a child. Additionally, the robot includes a speech synthesiser unit and is able to produce short spoken phrases using a neutral intonation.

The fourth category includes ball-shaped robots. We decided not to assign them to the mobile robots category, but to separate them due to their increased possibility of tactile interaction (i.e., kicking, picking up and holding). Like mobile robots, these also have additional functions besides movement. The Sphero uses built-in LEDs to have the effect of flashing and fading multicolored lights. Both robots have the function of emitting sound. Sphero plays music and Roball has the ability to communicate by vocal messages using a single chip device for voice recording and playback.

The last category contains other robots that have not been assigned to the described categories. The RBB robot has an arm with a basketball hoop attached to it. It can move in all directions (x, y and z axis) at different speeds with soft background music. The TWC looks like a soft, animal-shaped pillow. This robot was not assigned to the animal / creature category for the same reason that the ball-shaped robots were not assigned to the mobile robot category - more physical interaction (it is a wearable robot). TWC has four paws so that the child can wear it around the neck. There are sensors in the paws that detect the touch of the child and RGB LED strips that can illuminate them with different colors. There are speakers in the animal's head through which it is possible for a child to listen to brief sounds or music.

Table 4. Types of robots

Robot type	**Robot name**	**Articles**
humanoid	NAO	T2, T8, T9, T14, T16, T18, T25, T27, T37, T41, T47, T50, T73, S3, S6, S8
	Darwin-Mini (from Robotis)	T13, T17, T25, T30, T31, T34
	ZECA (Zeno Engaging Children with Autism)	T40, T42, T49, S7
	Kaspar	T19, T28, T54

	iRobiQ	T23, T26, T32
	CARO	T23, T26, T32
	FACE (an android face, developed by Hanson Robotics)	T54, T58
	Autonomous social robot	T72
	Bandit	T54
	Infanoid	T54
	Lego Mindstorms NXT	T53
	R-50 Alice	S4
	Robota	T54
	Tito	T54
animal / creature	Keepon	T21, T54, T68
	dinosaur robot Pleo	T36, T54
	Probo	T51, T55
	Muu	T54
	PABI	T48
	SAM	T20
	Semi-autonomous KiliRo robot	T24
	Zoomer dog	T6
mobile robot	ROMO	T13, T17, T30, T31, T39, S1
	Bubble blower	T54
	GIPY-1	T56
	Labo-1	S2
	Rovio (WowWee)	T37
	Remote-controlled robot	T5
ball-shaped robot	Sphero	T77
	Roball	T54
other	a new concept of social robot, called transitional wearable companion – TWC	T29
	Robot-based basketball (RBB)	T63

Taking into account the total number of articles where robots appeared, humanoid robots were the most popular (44). In second place were animals / creatures (12) and mobile robots (11). The popularity of individual robots within the given categories can be seen in Table 4. Considering only the articles describing the studies with participants without including articles reporting the general feasibilities of robots, the distribution is very similar, which can be seen in Table 5 showing the values also in percentage.

Table 5. Total number of articles where the given robots appeared

Robot type	**All articles**	**Only articles describing the studies with participants**
humanoid	44 (61.97 %)	32 (66.67 %)
animal / creature	12 (16.9 %)	8 (16.67 %)
mobile robot	11 (15.49 %)	6 (12.5 %)
ball-shaped robot	2 (2.82 %)	1 (2.08 %)
other	2 (2.82 %)	1 (2.08 %)
SUM	71	48

3.2. Robots over the years

In order to ensure the best readability of the charts (Figure 3), the presence of robots in the articles over the years was prepared for types of robots without division into individual robots, because there are 32 of them. Data labels have been added only for the "animal / creature" on one chart and "ball-shaped robot" on the second due to the data overlap in the charts. The rest of the results can be read easily without labels. An interesting relationship can be seen in the charts. In 2012, seven of the eight articles that featured humanoid robots were theoretical. Since 2014, the number of articles describing research with the use of humanoid robots has been increasing with each subsequent year until 2018. Mobile and animal-like robots have been used a similar number of times over the years. However, it can be seen that since 2017, interest in these types of robots has increased. Data for 2020 should be treated with reserve due to the fact that the literature search for this technical report ended in April 2020. Therefore, this report most probably does not cover all publications for 2020. Interestingly, in 2019, in contrast to previous years, humanoid robots were used almost as often as mobile and animal-like robots.

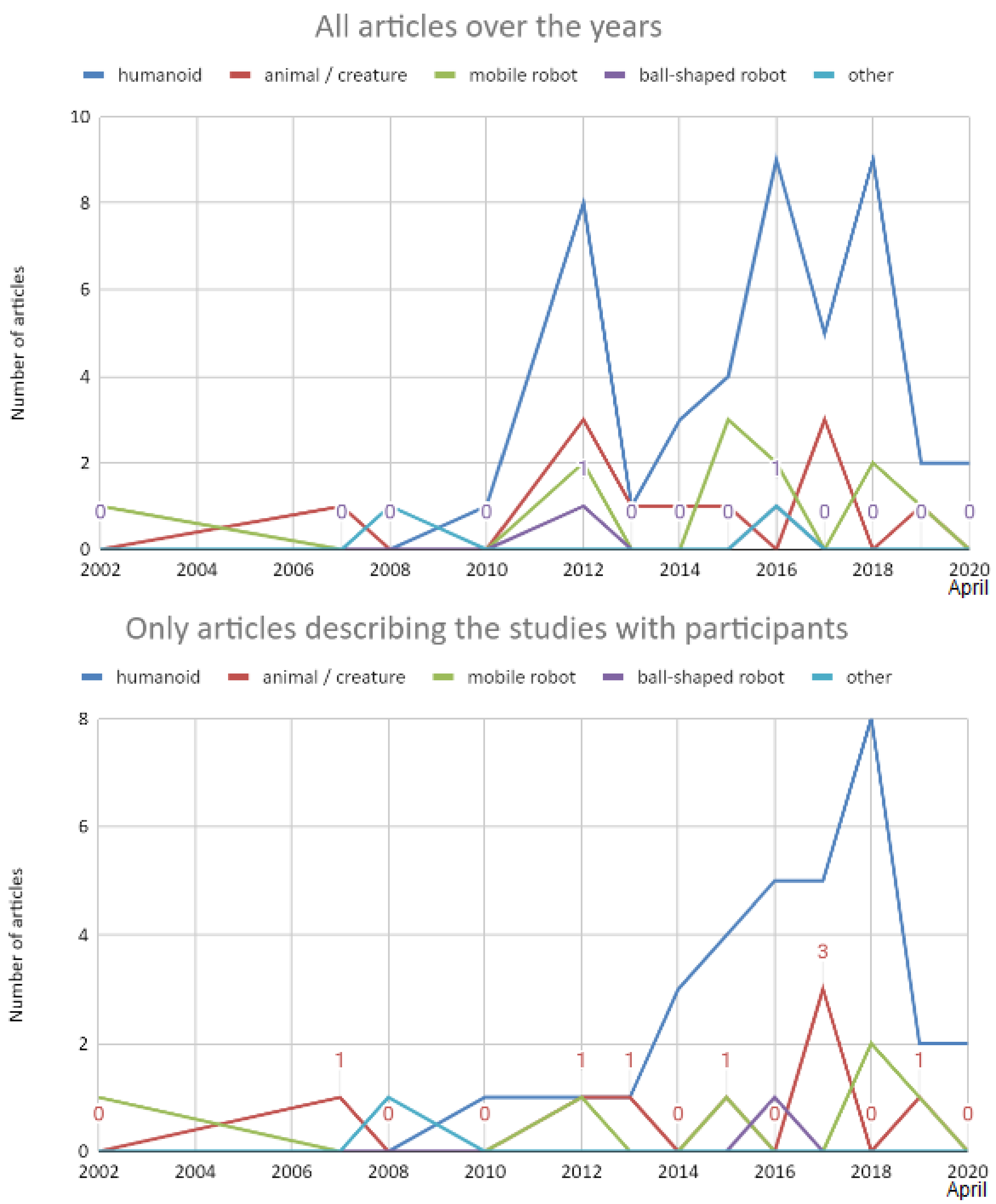

Figure 3. Number of all articles and only those describing research with the use of robots, broken down by types of robots

3.3. Robots used for age groups

We assigned each of the 30 studies that reported exact ages or age ranges of the participants to one of three age groups: (1) preschool age: 2-5 years, (2) primary school age: 6-8 years, (3) secondary school age: 10-14 years. Due to the large age ranges in some studies we decided to make the group assignments based on the youngest included child of a study. The 'preschool age group' included 17 studies, the 'primary school age' group included nine studies, and the 'secondary school age' group included four studies. As shown in Table 6, the robot NAO was used by eight studies from all age groups with the highest number of studies in the preschool age group. iRobiQ and CARO were used by three studies and Probo by two studies, but these robots were only used by studies assigned to the preschool age group. Keepon was used in the preschool age group (T68) and early school age group (T21). ZECA was only used in school age (T42, T49, S7).

Table 6. Articles assigned to three age groups based on the age of the youngest included participant. The third column depicts the robots that were used in the respective studies.

Age group	**Article**	**Robot used**
Preschool age (2-5 years)	T9, T16, T37, S6, S8 T23, T26, T32 T51, T55 T5 T6 T13 T19 T20 T36 T37 T68	NAO iRobiQ, CARO Probo Remote-controlled robot Zoomer Dog Darwin-Mini, ROMO Kaspar SAM Pleo Rovio Keepon
Primary school age (6-8 years)	T14, T50 T21 T24 T42 T53 T56 T58 T72	NAO Keepon KiliRo ZECA Lego Mindstorms NXT GIPY-1 FACE Autonomous social robot
Secondary school age (10-14 years)	T49, S7 T8 T63	ZECA NAO Robot-based basketball

4. Intervention

We analysed which skills were taught in robot-based interventions and grouped these skills in eight skill groups. Take into account that we only included studies with participants here; articles reporting the general feasibility of a robot to teach certain skills were not included. Table 7 shows the skill groups and skills taught in the respective robot-based intervention studies. The skill group taught by the highest amount of studies was 'social interaction: general' with 18 studies teaching related skills. The skills most often taught were 'recognition of basic emotions' by 15 studies and 'getting into attention' by ten studies.

Table 7. Skill groups and skills that were taught in the respective robot-based intervention studies. Some articles are listed twice or more often because they focused on two or more skills.

Skill group	**Skills**	**Articles**	**Robot used**
Social convention	Greeting skill	T6, T19, T28, T37, T41, T55, T73	Zoomer dog, Kaspar, NAO, Rovio, Probo

	Singing (children should learn to raise their hand when they want to do sth.)	T19, T41, T73	Kaspar, NAO
	Sharing	T6, T55	Zoomer dog, Probo
	Social convention: closing	T6	Zoomer dog
	Ability to thank	T55	Probo
	Basic speech: saying please	T73	NAO
Social interaction: General	Getting into interaction	T17, T19, T24, T26, T28, T36, T68, T72, T77, S2	Darwin-Mini, ROMO, Kaspar, KiliRo, iRobiQ, CARO, Pleo, Keepon, Autonomous social robot, Sphero, Labo-1
	Turn-taking	T6, T17, T28, T37, T72, S2, S6	Zoomer dog, Darwin-Mini, ROMO, Kaspar, NAO, Rovio, Autonomous social robot, Labo-1
	Following / imitating movements	T19, T58, S2	Kaspar, FACE, Labo-1
	Social attention abilities	T2, S6	NAO
	Social interaction	T24, T36	KiliRo, Pleo
	Child's engagement in an activity	T13, T20	Darwin-Mini, ROMO, SAM
	Social abilities	T72	Autonomous social robot
	Sensory processing skills	T13	Darwin-Mini, ROMO
	Basic eye contact	T32	iRobiQ, CARO
	Conversational interaction	T36	Pleo
	Socio-emotional behaviors	T13	Darwin-Mini, ROMO
Social interaction: Initiating	Focus on self-initiated interaction	T17	Darwin-Mini, ROMO
	Playing a puzzle game (child needs to ask robot for help)	T16	NAO
	Showing asked gestures (child needs to use a gesture to request for an object)	T53	Lego Mindstorms NXT
Social interaction:	Eye gaze following	T6, T23	Zoomer dog, iRobiQ, CARO

Responding	Response to a behavioral request	T6, S7	Zoomer dog, ZECA
	Response to name	T6	Zoomer dog
Emotions: Recognition	Recognition of basic emotions	T9, T14, T17, T19, T21, T23, T26, T42, T47, T49, T50, T51, S4, S7, S8	NAO, Darwin-Mini, ROMO, Kaspar, Keepon, iRobiQ, CARO, ZECA, Probo, R-50 Alice
	Sound making (mapping emotions and sounds)	T68, T77	Keepon, Sphero
	Context-emotion association	T9, S8	NAO
	Discrimination between thoughts and emotions	T9	NAO
	Reading emotions	T32	iRobiQ,CARO
Emotions: Expression	Mimic emotions learned in the robot training	T8, T14, S4, S7, S8	NAO, R-50 Alice, ZECA
Emotions: Control	Touching to transfer the robot to a positive emotional state	T77, T19	Sphero, Kaspar
	Adaptive behaviors in situations associated with anger and sadness	T21	Keepon
	Control negative emotions in social situations	T21	Keepon
Other skills	Self-care (self cleaning)	T41, T73	NAO,
	Cognitive skills	T53	Lego Mindstorms NXT
	Improve learning	T24	KiliRo
	Drumming game	T37	NAO, Rovio
	Rhythmic upper and lower body interpersonal synchrony	T37	NAO, Rovio
	Moving on a count	T37	NAO, Rovio
	Moving on a steady beat	T37	NAO, Rovio
	Selecting particular colors	T53	Lego Mindstorms NXT

We analysed the skill groups that were taught by the 30 studies we previously assigned to the three age groups ‘preschool age’ (2-5 years; 17 studies), ‘primary school age’ (6-8 years; 9 studies), and ‘secondary school age’ (10-14 years; 4 studies). As shown in Figure 4, the most frequently taught skill group of the ‘preschool age’ studies was ‘social interaction: general’, followed by ‘emotions: recognition’. Both skill groups were also the most frequently taught in the ‘primary school age’ intervention studies; however, in this age group ‘emotions: recognition’ was taught by more studies than ‘social interaction: general’. Three of the four studies of the ‘secondary school age’ group taught ‘emotions: recognition’, but none of them taught ‘social interaction: general’. ‘Social convention’ was only taught in ‘preschool age’ studies. ‘Emotions: expression’ was taught by two of the four ‘secondary school age’ studies, but only by a small proportion of the ‘preschool age’ and the ‘primary school age’ studies.

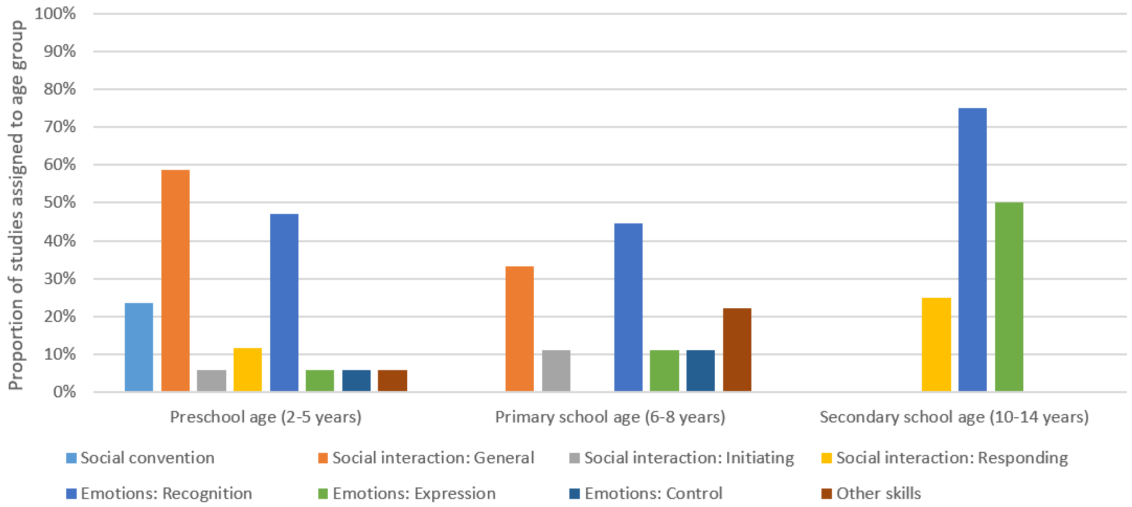

Figure 4. Proportion of studies per age group teaching skills of the respective skill groups.

Figure 4 shows the number of studies using the respective robots for interventions related to the eight skill groups. The robot NAO was used for the most skill groups, namely 6/8. NAO was the most frequently used robot for interventions on ‘social convention’, ‘social interaction: general’, ‘emotions: recognition’, ‘emotions: expression’, and ‘other skills’.

4.1. Emotion recognition

We focused on studies with participants and did not include here articles reporting the general feasibility of a robot to teach emotions. Nineteen studies provided intervention regarding emotion recognition. We provide the taught emotions in Table 8. We subsumed ‘happiness’ and ‘joy’, ‘fear’ and ‘scared’, and ‘anger’ and ‘annoyed’, respectively. This results in a total of seventeen different emotions including the neutral condition that were taught to be recognised (Table 8). T26 did not specify the taught emotions. The recognition of the emotions ‘happiness/joy’ and ‘sadness’ was most often taught. The recognition of ‘curious’, ‘proud’, ‘pleased’, ‘frustrated’, and ‘nervous’ were only taught by one study each.

Table 8: Emotion recognition taught in robot-based interventions. The third column shows the applied robots.

Emotion recognition: Emotion taught	**Article**	**Robot used**
Happiness/joy	T8, T9, T47, T50, S8 T42, T49, S7 T23, T32	NAO ZECA iRobiQ, CARO

	T51 T68 T77 T17 T19 S4	Probo Keepon Sphero ROMO, Darwin-Mini Kaspar R-50 Alice
Sadness	T8, T9, T47, T50, S8 T42, T49, S7 T23, T32 T51 T77 T17 T19 T21 S4	NAO ZECA iRobiQ, CARO Probo Sphero ROMO, Darwin-Mini Kaspar Keepon R-50 Alice
Fear/scared	T8, T9, T14, T47, T50, S8 T42, T49, S7 T68 T77 T17 T19 S4	NAO ZECA Keepon Sphero ROMO, Darwin-Mini Kaspar R-50 Alice
Anger/annoyed	T8, T9, T14, T47, T50, S8 T42, T49, S7 T23, T32 T77 T17 T21	NAO ZECA iRobiQ, CARO Sphero ROMO, Darwin-Mini Keepon
Surprised	T42, T49, S7 T23, T32 T17 S4	ZECA iRobiQ, CARO ROMO, Darwin-Mini R-50 Alice
Disgust	T8, T9, T47 T42 T17 S4	NAO ZECA ROMO, Darwin-Mini R-50 Alice
Shy	T23, T32 T47 T17	iRobiQ, CARO NAO ROMO, Darwin-Mini
Neutral condition	T8 T42 S4	NAO ZECA R-50 Alice
Hungry	T14, T47, T50	NAO
Excited	T68 T17	Keepon ROMO, Darwin-Mini
Tired	T47	NAO

	T17	ROMO, Darwin-Mini
Love/hug	T47, T50	NAO
Curious	T17	ROMO, Darwin-Mini
Proud	T17	ROMO, Darwin-Mini
Pleased	T17	ROMO, Darwin-Mini
Frustrated	T17	ROMO, Darwin-Mini
Nervous	T17	ROMO, Darwin-Mini
Emotions not specified	T26	iRobiQ, CARO

Figure 5 shows all robots that were used for emotion recognition interventions. The most different emotions were taught by the robots ROMO and Darwin-Mini, followed by NAO (Figure 5). The robot Probo was used to teach only two emotions.

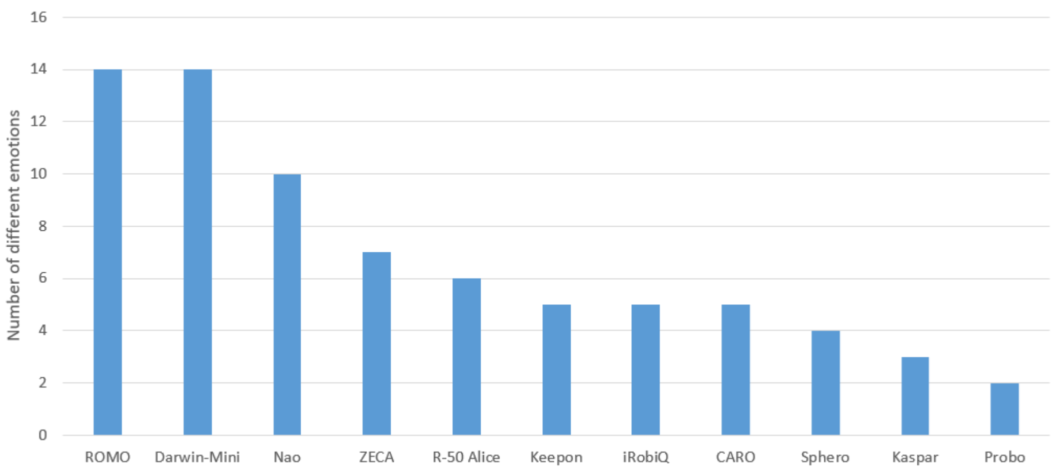

Figure 5: Total number of different emotions taught by each robot over all articles of the SLR.

4.2. Emotion expression

Five studies taught the children to produce/imitate specific gestures that express emotions. We provide the taught emotions in Table 9. We subsumed 'happiness' and 'joy', 'fear' and 'scared', and 'anger' and 'annoyed', respectively. This results in a total of eight different emotions including the neutral condition that were taught to be expressed (Table 9). The expression of 'fear/scared' was most often taught, namely by all five studies. T14 was the only study to teach the expression of 'hungry'. Three different robots were used to teach the expression of emotions: NAO was used to teach the expression of 7/8 emotions, R-50 Alice was used for 6/8 emotions, ZECA was used for 5/8 emotions.

Table 9: Emotion expression taught in robot-based interventions. The third column shows the applied robots.

Emotion expression: Emotion taught	Article	Robot used
Fear/scared	T8, T14, S8 S4 S7	NAO R-50 Alice ZECA
Happiness/joy	T8, S8 S4 S7	NAO R-50 Alice ZECA
Anger/annoyed	T8, T14, S8 S7	NAO ZECA
Sadness	T8, S8 S4 S7	NAO R-50 Alice ZECA
Disgust	T8 S4	NAO R-50 Alice
Surprised	S4 S7	R-50 Alice ZECA
Neutral condition	T8 S4	NAO R-50 Alice
Hungry	T14	NAO

4.3. Emotion control

Three studies taught the children to control emotions. T19 teached the participants to make the robot happy again when it is sad. T21 used the robot Keepon to teach the control of sadness and anger. T77 used the robot Sphero to teach the participants to transfer the robot to a positive emotional state when touching it.

5. Challenges

We investigated the faced challenges, limitations, and concerns reported by the articles.

5.1. Robot-child interaction

Eight articles reported challenges, limitations, and concerns related to the interaction of children with the robot during intervention sessions (Table 10).

Table 10: Reported challenges, limitations, and concerns related to robot-child interaction. The third column provides the names of the robot(s) that were used in the respective articles.

Challenge/Limitation/Concern	Article	Robot used
Concerns that interacting with a robot may be too complex for a child without the help of a therapist	T26	iRobiQ, CARO

Only one-on-one interaction possible	T2	NAO
Distractors on robot	T2	NAO
No/too little adaptation to therapeutic needs of children	S3	NAO
Challenge to determine the responses of child to training stimuli	T32	iRobiQ, CARO
Personalizing automated engagement estimation to each child with autism	S8	NAO
Interaction of children with robot	T42 S4	ZECA R-50 Alice
Unstructured and unconstrained interactions	S2	Labo-1
Gender of robot affects interaction	S4	R-50 Alice
No multimodal interaction used	S4	R-50 Alice

5.2. Intervention success

Three articles discussed challenges, limitations, and concerns related to the success of robot-based interventions (Table 11).

Table 11: Reported challenges, limitations, and concerns related to intervention success. The third column provides the names of the robot(s) that were used in the respective articles.

Challenge/Limitation/Concern	**Article**	**Robot used**
Emotions cannot be recognised/distinguished well by participants	T25	NAO, Darwin-Mini
Concerns about intervention success in cases such as a child-centered approach to teaching or unilateral verbal expression towards the robot	T26	iRobiQ, CARO
It remains open whether children understood animations in intervention phase	T51	Probo

5.3. Target group

Six articles reported challenges, limitations, and concerns regarding the target group of children with ASD (Table 12).

Table 12: Reported challenges, limitations, and concerns related to the target group. The third column provides the names of the robot(s) that were used in the respective articles.

Challenge/Limitation/Concern	Article	Robot used
Low expressivity of children	T73	NAO
Variability on level of functioning of the children	T73	NAO
Variability on the compliance levels of parents with training	T37	NAO, Rovio
Limited attention span of children	S3	NAO
Low expressivity of children	T63 T73	Robot-based basketball NAO
Tasks are too complex for some children	T53 S2	Lego Mindstorms NXT Labo-1
Wearing physiological sensors might be problematic for some children	T63	Robot-based basketball

The articles reporting low expressivity of children (T63, T73) as a challenge included both low functioning and high functioning (T63) or only high functioning participants (T73) with ASD. The articles reporting that tasks were too complex for some children (T53, S2) included children with ASD without further assignment to ASD subgroups.

5.4. Methodological issues

Nineteen articles discussed the small sample size and the related limited generalisability of their findings and three articles discussed limitations related to the heterogeneous sample of participants. Table 13 depicts the number of participants with ASD and the ASD subgroups for the respective articles. The 19 articles reporting a small sample size included 1-27 participants with ASD, respectively. The mean number of participants with ASD was 9.58, the median was 9, and the standard deviation was 6.84.

Table 13: Articles reporting challenges, limitations, and concerns related to small sample size and heterogeneous sample, respective number of participants with ASD and included ASD subgroups. Abbreviations: ASD = autism spectrum disorder; ASD-gen: Participants with ASD without further assignment to subgroups; HF = high functioning; LF = low functioning.

Challenge/Limitation/Concern	Article	Participants with ASD	ASD subgroup
Small sample size/limited generalisability of findings	T2	11	ASD-gen
	T5	21	ASD-gen

	T8	1	ASD-gen
	T9	14	ASD-gen
	T13	3	1 LF, 2 HF
	T14	13	LF
	T16	8	ASD-gen
	T19	2	LF
	T20	13	HF
	T21	27	ASD-gen
	T23	15	ASD-gen
	T24	9	ASD-gen
	T41	2	ASD-gen
	T51	3	ASD-gen
	T58	5	ASD-gen
	T77	12	ASD-gen
	S2	6	ASD-gen
	S4	14	6 LF, 8 HF
	S7	3	HF
Heterogeneous sample	T53	14	ASD-gen
	S2	6	ASD-gen
	S4	14	6 LF, 8 HF

Three studies (T5, T9, T23) raised limitations related to gender ratio. T5 included 20 males and one female with ASD, T9 included 12 males and two females with ASD, and T23 included 15 males and no female with ASD. T58 and T73 reported challenges regarding the combination of robots and emotion recognition. T58 furthermore reported challenges regarding the integration of data from different sensors. T63 discussed low interrater reliability in determining emotions of children and the challenge that several affective states could co-occur at different arousal levels. S2 discussed potential interviewer bias and S8 discussed the potential influence of the annotators' cultural background on their annotations. Eleven studies (T6, T9, T13, T18, T20, T21, T23, T24, T37, T51, S4) discussed further limitations related to study design and procedures (e.g., problems related to the room where the experiments were carried out, order of appearance of robots may have influenced results, additional experiments or outcome measures are warranted, limitations of applied interviews and questionnaires).

6. Recommendations

6.1. Longitudinal studies

In four articles (T26, T32, S8, T54), researchers suggested that longitudinal studies over an extended time period should be conducted in the future. In two studies (T26, T32) each child attended eight clinical sessions (each lasted approximately 30–40 min and contained ten trials for two training interactions). The third study (S8) focused on single day recordings of the children. The fourth article (T54) is a review and indicates that described studies over a few days or, rarely, a few weeks or months.

6.2. Study design

In six articles (T5, T9, T14, T21, T23, T37), authors report that additional measures/assessments should be added in future studies. In two articles, researchers indicate the need to conduct research also in a different environment, e.g., school or group sessions (T21), research within clinical facilities, but also in the child's natural environment (T16).

6.3. Course of the intervention sessions

Recommendations regarding the course of the intervention sessions concern interruptions of the sessions, verbal instructions given to the child and objects used during the interventions. T2 suggested that sessions should be conducted without interruptions. T34 recommended the wording of the verbal instructions to be short, brief, simple, and concrete. T37 recommended the inclusion of object-free, creative movement interventions involving rhythm, dance, yoga, and play therapies into the standard-of-care treatment of children with ASD.

6.4. Participants

This section mainly describes the recommendations on the warranted characteristics of children participating in studies on robot-based interventions. Researchers indicate that certain motor skills (T49) and some verbal response capacities (T2) of children are necessary. Some studies (T6, T32, S7) recommended to include children with ASD with different levels of functioning. In addition, studies should include a control group of typically developing children (T23, S7). The articles also included recommendations regarding children's experience with robots or a given type of therapy. Regarding the children's experience with a robot, opinions are divided: T16 recommended that children should not have previous experience with a robot, so as to rule out the familiarity effect (the participants of that study had previous experience with the robot). However, T53 even recommended a familiarisation phase, where the children can freely explore the robot in order to reduce anxiety levels and make the robot more attractive. In case of Pivotal Response Treatment (PRT) therapy, testing participants who have no or limited experience with this type of therapy to minimize the floor effect observed in the average percentage of required prompts, is indicated (T16). The final two recommendations concern the course of the session and the reporting of the results: Researchers recommended to increase the number of trials per participant to reduce between-participant variance (T16). When describing the results, characteristics of the participant group should be provided (T9).

6.5. Engaging of children

In one article, researchers suggest that future research should develop diverse training activities that can sustain children's engagement over prolonged training durations (T37). However, in several others, researchers suggest not only ways to engage children, but also possibilities how to measure engagement. The engagement-related suggestions apply both to robots and the reward systems. According to T14, a real social robot is more engaging than an animated one. Another study (T63) found that affect-adaptive gameplay with a robot was more engaging and showed more liking in comparison with non-adaptive gameplay. Regarding rewards, in one study, a reinforcement (snacks or access to toys) was offered by the teacher at the end of each pretest, post test and training session (T14). Other researchers suggest reward systems should be more elaborate and customized to individual subjects' needs, interests, and social abilities (T23). As measurement tools, researchers propose Modified Fogg's Behavioral Model (MFBM) to demonstrate the motivation, i.e., level of engagement shown by the children while interacting with a robot (T18, T73). In another study researchers presented initial design schemes of the robotic framework for initiating and estimation engagement (S1). A musical stimulus was be used for initiating engagement. The system needs to be able to detect emotional and social states of a child. Once perceived, it is imperative that the robotic system displays appropriate expressive behaviors and stimulating motions to engage a child emotionally and socially. Researchers use RGB-D depth sensors (e.g. Microsoft Kinect) to monitor the physical activities of a child to estimate the social engagement.

6.6. Robots - NAO, Darwin-Mini, ZECA, Probo, Pleo

The articles collected experiences with the use of individual robots, which may be a guide for future research applying robots for interventional purposes. In one of the studies (T25), a system to teach five of the six universal emotions was developed, i.e., sadness, anger, happiness, fear, and surprise (disgust was not investigated). The researchers created an emotional gesture set that has been performed on the NAO and the Darwin-Mini robots. In order to find out how well the set of gestures can be recognised, they conducted a pilot study with able-bodied adults. Participants were supposed to observe the gestures produced by robots and determine what emotion they were attempting to portray. The relatively simple robot (Darwin-Mini) was able to express happiness better while the more advanced robot (NAO) was able to express sadness better. The study participants had difficulty differentiating between some emotions expressed by the NAO robot: Happiness and surprise were commonly confused with one another, as were fear and sadness. NAO expresses happiness through speedily raising his arms above his head, forming a wide "V" with his arms and sways from side to side energetically. Surprise is represented by very abruptly holding both arms, shoulder length apart, above his head and back into the neutral pose after a few seconds. To represent sadness, NAO either crouches down towards the ground with one arm outstretched and the other hand shielding his face or crouches down towards the ground with both hands covering his face while shaking uncontrollably. To express sadness, NAO either bends his knees slightly, leans forwards, and brings his hands up toward his face while shaking his head slowly from side to side or bends his knees, and slowly shakes his head in the crook of his bended arm while the other arm is flat at his side. Therefore, when creating future gesture sets, the focus should be on a better differentiation between these emotion sets.

In another study, using the Darwin-Mini robot (T34), researchers also make recommendations on the gestures that represent emotions performed by the robot. Happiness and Sadness got higher scores because of the clear design features. For Happiness, the robot body is open, positive and free with hands waving and eyes looking up. For Sadness, the robot is bent down with banging its head against its hands and touching its eyes. However, there are some features making them confusing, such as the fast moving arms which make people also feel the robot is angry. For Anger, the arms are in a defensive position and eyes are looking into the air, while a few people feel it is bragging. Last, covering the eyes, defending itself makes it fear. However, shielding its face makes people feel it is surprised. High recognition rates were achieved, but more efforts are needed to adjust the robot body expressions in such a way that they would be perceived consistently among participants. Moreover, the emotion expressed is not exactly like that in real life, since expression in real life is subtler. To express emotion more authentically, the motions should be smaller.

Another study using the NAO robot (T41) adopted Kansei Engineering due to its unique mechanism in extracting implicit feeling and emotion. Kansei Engineering is a product development methodology that translates customers' and users' feelings, impressions, and emotions into concrete design parameters [4]. There are several ways to measure Kansei: words, psychological reaction (heart rate, EMG, EEG), behavior, facial and body expression. Kansei words are determined by the field. They are collected from all possible sources where words are used to describe the product field. Often included are: magazines, literature, manuals, experts, experienced users, ideas, imaginations [5]. Researchers collected the Kansei words related to humanoid-robot interaction and children with ASD from past literature involving the use of robots for intervention and then arranged into a checklist form using five point semantic differential scales. Examples of words include: attractive, depressing, and stimulating. In the instances where a child is unable to fill in the Kansei checklist, the teacher observes and fills in the checklist while the child is involved in the interaction. Results showed the highest score for almost all the positive emotions in the module "Let's Sing a Song!" when the robot NAO was singing and dancing. Additionally, researchers noted on the video footage review that the children were showing a more apparent emotional response while the robot talked or made hand gestures.

The study that used the ZECA robot (called the Zeno R-50 there) (T42) consisted of eight stages. Stage 3, 5, and 7 were the main experiment game. For each game stage, a sequence of 13 emotions was randomly shown by Zeno. After showing each emotion, Zeno would resume a neutral pose and wait for the child's response. The child's task was to predict/guess what emotion the robot is trying to show. It was a comparative study testing how children diagnosed with ASD vs. TD children can recognize emotion expressions as shown by a humanoid robot. Researchers also studied the effect of using body gestures by the robot on the expression recognition predictions of children. In total, 37 animations were used. Nineteen of these consisted of only facial expressions, including a neutral expression. The other 18 animations were each based on the gesture mimicry of 18 human actors. The animations were developed to closely mimic the expressions of the actors' facial and upper body expressions with the available degrees of freedom, because researchers have not developed software for Zeno's lower body kinematics at the moment. Addition of gestures for Happy greatly lowered the guess accuracy in both groups. This is due to the fact that the gestures could represent more than one emotion, e.g., happiness and surprise when meeting an unexpected friend. However, a significant improvement was shown for Disgust in the ASD group but not TD when gestures were added. Since the robot lacks the ability to show nose wrinkler, it is understandable why Disgust had a low recognition rate for facial expression. Also, a significant improvement was shown in Fear recognition for the TD group but not the ASD group. The TD group was able to greatly benefit from the incorporation of gestures for Fear by incrementing from a 12% to a 66% accuracy. Although the ASD group slightly improved with the addition of gestures for Fear, the increment is not comparable to that found by the TD group. The improvement difference was so dramatic that this proved to be the only skill significantly differentiating the ASD vs TD groups. Through these findings they demonstrated that a general impairment in expression recognition for children with ASD should not be assumed when designing SAR (Socially Assistive Robotics) therapies for them. Instead, each emotion should be evaluated individually. The researchers showed that the use of gestures can significantly impact the prediction accuracy of both children with ASD and TD children in a negative or positive manner depending on the emotion. Although the children with ASD did not show any significant impairment for correctly labeling most expressions, future work should investigate whether the children can truly identify the emotional meaning connected to the label and visual cue. To test this, for example the child can be asked to make up a short story explaining why the robot may be showing such emotion expression. As this may prove difficult for some children to do, the use of electro dermal activity to identify if the child associates the expression with an emotion can also be explored.

One of the studies using the Probo robot (T51) had two phases, one phase consisted of watching a video that was played on Probo's belly representing a situation that generates an emotion followed by a neutral facial expression of the robot (Phase A: exposure to Probo's neutral face), and the other phase was identical to the first one, with the difference that the video was followed by facial expression of the robot with the right emotion (Phase B: exposure to Probo's active face). The emotion recognition performance was recorded for each participant in two phases. In each exposure, the participant had to recognize one of the two emotions, i.e., happiness or sadness, from an animation in which something positive or negative was happening with Probo. To assess their performance in each task, a binary scale was used, i.e., correct or wrong recognition of the emotion suggested by each animation. The results show that the performance of participants improved when Probo's active face was used, compared with the phase where Probo expressed a neutral face. Additionally, using Probo's active face has similar effects in increasing the emotion recognition performance of both happiness and sadness.

In the second study using the Probo robot (T55), the main experimenter together with the child's therapist and parents identified a specific social skill deficit. Then, they selected the following target skills: "sharing toys" (for two participants), saying "thank you" and saying "hello". An individualized Social Story was developed for each of the skills. In one phase of the study, Probo was telling a story. The robot also expressed the emotions that were included in the story, i.e., happiness and sadness, and moved its head, eyes and trunk. After the robot had told the story, the child had to exercise the social ability described above that was targeted in the story. The story is played on the robot without interruption, so the therapist cannot stop the story when necessary. Accordingly, researchers recommend introducing more interactive stories, so that the robot can respond to the actions and the reactions of a child during the presentation of the social stories. Moreover, the two therapists involved in this study offered valuable feedback on the design of the robot. The size of the robot (80 cm) appears to be appropriate for interaction with children as is the relative size of the head compared to the body so it is easy for the children to focus their attention on the facial expressions. The face area (eyes and mouth) is not really a triangle, but rather a rectangle, which probably needs more attention resources than a triangle does. Also, it was suggested that a shorter trunk would increase the visibility of the facial expressions of Probo, so the mouth would be more visible to the children when the trunk is in the down position. Therapists did not make any remarks or negative reactions to the green color of Probo's coat. A demonstration was held for the other children of the autism center and the researchers observed that even those children who had problems with touching and being touched, were able to touch and interact with the robot at the end of the session.

In another study (T36), the researchers used the commercially marketed, toy dinosaur robot Pleo they modified to play pre-recorded, synchronized motor and audio behaviors, triggered through the robot's infrared receiver by a television remote control. The robot continuously performed varying waiting behaviors (termed 'idling' in the paper), shifting its body weight, looking around the room, occasionally muttering or humming to itself. In order to instill a belief that the robot was capable of rich social interaction, the robot expressed positive (e.g., excitement and joy) and negative (disappointment or dislike) affect and attention to particular objects using non-linguistic vocalizations (e.g., "Ooooohh!" to indicate interest in an object, and, "Awwww" to indicate disappointment), head turns, and/or body movements. The researchers chose non-linguistic vocalisations instead of speech because they lacked an interface for spontaneous unplanned speech production. Therefore, they designed the robot to preemptively avoid children asking the robot to express thoughts beyond its affective or attentional state. To make the interaction more enjoyable, they chose a cartoonish vocal character, designed a happy dance to express elation, and offered participants a chance to pet the robot. Researchers studied only children with high functioning ASD, primarily because they expected that lower functioning children who, by definition, have greater difficulty with language and lower cognitive functioning, would be more difficult to restrict to largely non-tactile interaction with the robot. For younger or lower-functioning children, they suggest that more robust robotic platforms should be used.

The last two paragraphs refer to theoretical articles. One article (S3) presents the proposed experimental layout of a preliminary test to observe the initial response and behavior of children with ASD when they are exposed to the humanoid robot NAO. The robot can be controlled by the program using Windows, MAC OS, and even Linux. The authors presented the tool Choregraphe that allows even complicated behavior to be programmed easily through its user-friendly GUI. This is valuable, because non-technical experts will be able to use NAO to execute pre-programmed behaviors using this software. Choregraphe also allows users to create new, predefined motions using the Python programming language.

Nine years ago, the authors of a review article regarding robots for use in research on ASD (T54) described that before robots can become autonomous units in therapeutic interactions, beyond detecting and responding to user actions, they must be able to sense the moods and preferences of users and adjust their behavior in real time to these factors. At that time, the authors point out that some systems that identify mental and emotional states from physiological and behavioral data have already been proposed, but more work in this area is needed before robots can reliably respond to users' moods and preferences.

6.7. Learning gestures

The researchers of T14 developed a robot-based intervention in which a social robot, NAO, taught children with ASD how to recognize (Phase I) and produce (Phase II) eight pantomime gestures that express feelings and needs that are commonly used in daily life. It is crucial to teach children with ASD the meanings of emotional gestures before asking them to imitate the gestures. Gestures had consistency rates of 70% or above. Each gesture lasted for three to four seconds. The robot gestured while narrating scenarios. Besides verbal narration, the scenarios were visually displayed on the laptop screen, which was placed next to NAO. Two pictures were drawn for each scenario and they were presented sequentially as NAO narrated the corresponding scenes. The intervention program lasted for 12 weeks (with each phase lasting for six weeks). Four sets of narrations (S1, S2, S3, S4), each set containing eight different scenarios describing the use of the eight gestures, were designed for each phase. One of its scenarios for an ANGRY gesture was: "Sister is playing with a toy car with her brother. She breaks it. Brother feels angry and he does this". Each scenario contained two to three sentences. The four scenarios for the same gesture shared a similar length but contained different contents. Each phase consisted of four pre-tests, one for each set of narrations (S1, S2, S3, S4), four training sessions (two training sessions for S1 and S2, with two sessions per week), four immediate post-tests that were the same as the pre-tests, and the same follow-up post-tests after two weeks. The post-tests for S1 and S2 assessed the training effects, while those for S3 and S4 assessed the generalization effects. Each session lasted for approximately 30 min.

Results of the study suggest that the participants might not have been able to generalize the acquired gestural production skills when interacting with human beings. Results showed that participants performed more poorly in Phase II than in Phase I. Perhaps it is more difficult for children with ASD to master the gestural production skills than the gestural recognition skills. Gestural production requires gestural recognition as well as fine motor skills. Therefore, it might take more than four training sessions to equip the participants in the intervention group with the necessary gestural production skills. The authors suggested that future studies should provide the children with ASD with more training sessions regarding gestural production.

6.8. Learning outcomes

In study T14, described in the previous section 6.7., researchers concluded it should be investigated whether or not the learning outcomes can be maintained for a longer period of time (i.e., beyond two weeks). In order to examine this, they recommend observations of the behaviors of the participants in schools and at home for an extended period of time. In another study (T56), researchers came to similar conclusions. They said further experiments are needed to investigate the repeatability and durability of the effects. The study examined the role of a mobile robot named "GIPY-1" in the context of social and emotional interaction of a child with ASD with a third person, namely a therapist. The three-pronged interaction among the child with ASD, the robot and the therapist was investigated in spontaneous, free game play by means of a multimodal approach. The researchers analysed the interaction between the robot and the child using different criteria such as eye contact (looking at the robot), manipulation (operating with the robot), touch (touching the robot without manipulating it), and posture (changing postural position towards the robot). The criteria researchers have chosen represent the state of the child's cognitive processes, as expressed by the interest the child exhibits towards the robot in spontaneous, free game play. Once this state is established, the child develops a triadic relation with the robot and the therapist, thereby displaying enjoyment, which is a positive emotion. This expression appears when the child interacts with the therapist using the robot. Positive emotion is quasi absent when the child interacts with the mobile robot on a standalone basis.

Acknowledgements

This publication was supported by the European Commission's Erasmus+ project 'EMBOA, Affective loop in Socially Assistive Robotics as an intervention tool for children with autism', contract no 2019-1-PL01-KA203-065096.

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Included articles in the SLR

[T2] Alnajjar, F., Cappuccio, M., Renawi, A., Mubin, O., & Loo, C. K. (2020). Personalized robot interventions for autistic children: An automated methodology for attention assessment. *International Journal of Social Robotics*, 1-16.

[T5] Zantinge, G., van Rijn, S., Stockmann, L., & Swaab, H. (2019). Concordance between physiological arousal and emotion expression during fear in young children with autism spectrum disorders. *Autism*, *23*(3), 629-638.

[T6] Silva, K., Lima, M., Santos-Magalhães, A., Fafiães, C., & de Sousa, L. (2019). Living and Robotic Dogs as Elicitors of Social Communication Behavior and Regulated Emotional Responding in Individuals with Autism and Severe Language Delay: A Preliminary Comparative Study. *Anthrozoös*, *32*(1), 23-33.

[T8] Conti, D. Trubia, G. Buono, S. Di Nuovo, S. Di Nuovo, A. (2019). Affect Recognition in Autism: A single case study on integrating a humanoid robot in a standard therapy. *Qwerty,* 14(2), 66-87.

[T9] Marino, F. Chilà, P. Sfrazzetto, S.T. Carrozza, C. Crimi, I. Failla, C. Busà, M. Bernava, G. Tartarisco, G. Vagni, D. Ruta, L. Pioggia, G. (2019). Outcomes of a Robot-Assisted Social-Emotional Understanding Intervention for Young Children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders.*

[T13] Javed, H. Jeon, M. Park, C.H. (2018). Adaptive Framework for Emotional Engagement in Child-Robot Interactions for Autism Interventions. *15th International Conference on Ubiquitous Robots,* 396-400.

[T14] So, W.-C. Wong, M.K.-Y. Lam, C.K.-Y. Lam, W.-Y. Chui, A.T.-F. Lee, T.-L. Ng, H.-M. Chan, C.-H. Fok, D.C.-W. (2018). Using a social robot to teach gestural recognition and production in children with autism spectrum disorders. *Disability and Rehabilitation: Assistive Technology,* 13(6), 527-539.

[T15] Joseph, L. Pramod, S. Nair, L.S. (2018). Emotion recognition in a social robot for robot-assisted therapy to autistic treatment using deep learning. *2017 IEEE International Conference on Technological Advancements in Power and Energy, TAP Energy 2017*, 1-6.

[T16] van Straten, C.L. Smeekens, I. Barakova, E. Glennon, J. Buitelaar, J. Chen, A. (2018). Effects of robots' intonation and bodily appearance on robot-mediated communicative treatment outcomes for children with autism spectrum disorder. *Personal and Ubiquitous Computing*, 22(2), 379-390.

[T17] Javed, H. Jeon, M. Howard, A. Park, C.H. (2018). Robot-Assisted Socio-Emotional Intervention Framework for Children with Autism Spectrum Disorder. *13th Annual ACM/IEEE International Conference on Human Robot Interaction, HRI 2018,* 131-132.

[T18] Aziz, A.A. Moganan, F.F.M. Mokhsin, M. Sakamat, N. Ismail. (2018). A. Humanoid robot intervention with autism: The conceptual model in demonstrating the emotional responses of children with autism. *7th International Conference on Kansei Engineering and Emotion Research, KEER 2018,* 574-579.

[T19] Zorcec, T. Robins, B. Dautenhahn, K. (2018). Getting Engaged: Assisted Play with a Humanoid Robot Kaspar for Children with Severe Autism. *10th International ICT Innovations Conference, ICT Innovations 2018,* 198-207

[T20] Koch, S.A. Stevens, C.E. Clesi, C.D. Lebersfeld, J.B. Sellers, A.G. McNew, M.E. Biasini, F.J. Amthor, F.R. Hopkins, M.I. (2017). A Feasibility Study Evaluating the Emotionally Expressive Robot SAM. *International Journal of Social Robotics,* 9(4), 601-613.

[T21] Costescu, C.A. Vanderborght, B. David, D.O. (2017). Robot-Enhanced CBT for dysfunctional emotions in social situations for children with ASD. *Journal of Evidence-Based Psychotherapies*, 17(2), 119-132.

[T23] 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.

[T24] Bharatharaj, J. Huang, L. Mohan, R.E. Al-Jumaily, A. Krägeloh, C. (2017). Robot-assisted therapy for learning and social interaction of children with autism spectrum disorder. *Robotics*, 6(1).

[T25] English, B.A. Coates, A. Howard, A. (2017). Recognition of Gestural Behaviors Expressed by Humanoid Robotic Platforms for Teaching Affect Recognition to Children with Autism - A Healthy Subjects Pilot Study. *SOCIAL ROBOTICS, ICSR 2017,* Volume 10652 LNAI, 567-576.

[T26] Yun, S.-S. Choi, J. Park, S.-K. Bong, G.-Y. Yoo, H. (2016). Robotic behavioral intervention to facilitate eye contact and reading emotions of children with autism spectrum disorders. *25th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2016,* 694-699.

[T27] Hirokawa, M. Funahashi, A. Pan, Y. Itoh, Y. Suzuki, K. (2016). Design of a robotic agent that measures smile and facing behavior of children with Autism Spectrum Disorder. *25th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2016,* 843-848.

[T28] Huijnen, C.A.G.J. Lexis, M.A.S. de Witte, L.P. (2016). Matching Robot KASPAR to Autism Spectrum Disorder (ASD) Therapy and Educational Goals. *International Journal of Social Robotics*, 8(4), 445-455.

[T29] Özcan, B. Caligiore, D. Sperati, V. Moretta, T. Baldassarre, G. (2016). Transitional Wearable Companions: A Novel Concept of Soft Interactive Social Robots to Improve Social Skills in Children with Autism Spectrum Disorder. *International Journal of Social Robotics*, 8(4), 471-481.

[T30] Bevill, R. Azzi, P. Spadafora, M. Park, C.H. Kim, H.J. Lee, J.W. Raihan, K. Jeon, M. Howard, A.M. (2016). Multisensory robotic therapy to promote natural emotional interaction for children with ASD. *11th Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI 2016*, Volume 2016-April, 571-571.

[T31] Bevill, R. Park, C.H. Kim, H.J. Lee, J.W. Rennie, A. Jeon, M. Howard, A.M. (2016). Interactive robotic framework for multi-sensory therapy for children with autism spectrum disorder. *11th Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI 2016*, Volume 2016-April, 421-422.

[T32] Yun, S.-S. Kim, H. Choi, J. Park, S.-K. (2016). A robot-assisted behavioral intervention system for children with autism spectrum disorders. *Robotics and Autonomous Systems*, 76, 58-67.

[T34] Wang, H. Hsiao, P.-Y. Min, B.-C. (2016). Examine the potential of robots to teach autistic children emotional concepts: A preliminary study. *SOCIAL ROBOTICS, (ICSR 2016)*, Volume 9979 LNAI, 571-580.

[T36] Kim, E.S. Daniell, C.M. Makar, C. Elia, J. Scassellati, B. Shic, F. (2015). Potential clinical impact of positive affect in robot interactions for autism intervention. *2015 International Conference on Affective Computing and Intelligent Interaction, ACII 2015,* 8-13.

[T37] Srinivasan, S.M. Park, I.K. Neelly, L.B. Bhat, A.N. (2015). A comparison of the effects of rhythm and robotic interventions on repetitive behaviors and affective states of children with Autism Spectrum Disorder (ASD). *Research in Autism Spectrum Disorders*, 18, 51-63.

[T39] Zhang, R. Jeon, M. Park, C.H. Howard, A. (2015). Robotic Sonification for Promoting Emotional and Social Interactions of Children with ASD. *10th Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI 2015*, Volume 02-05-March-2015, 111-112.

[T40] Leo, M. Coco, M.D. Carcagnì, P. Distante, C. Bernava, M. Pioggia, G. Palestra, G. (2015). Automatic Emotion Recognition in Robot-Children Interaction for ASD Treatment. *15th IEEE International Conference on Computer Vision Workshops, ICCVW 2015*, Volume 2015-February, 537-545.

[T41] Aziz, A.A. Moganan, F.F.M. Ismail, A. Lokman, A.M. (2015). Autistic Children's Kansei Responses Towards Humanoid-Robot as Teaching Mediator. *IEEE International Symposium on Robotics and Intelligent Sensors, IEEE IRIS 2015*, 76, 488-493.

[T42] Salvador, M.J. Silver, S. Mahoor, M.H. (2015). An emotion recognition comparative study of autistic and typically-developing children using the zeno robot. *2015 IEEE International Conference on Robotics and Automation, ICRA 2015*, Volume 2015-June, 6128-6133.

[T47] Miskam, M.A. Shamsuddin, S. Samat, M.R.A. Yussof, H. Ainudin, H.A. Omar, A.R. (2014). Humanoid robot NAO as a teaching tool of emotion recognition for children with autism using the Android app. *2014 International Symposium on Micro-NanoMechatronics and Human Science, MHS 2014*.

[T48] Dickstein-Fischer, L. Fischer, G.S. (2014). Combining psychological and engineering approaches to utilizing social robots with children with Autism. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, 792-795.

[T49] Costa, S. Soares, F. Pereira, A.P. Santos, C. Hiolle, A. (2014). Building a game scenario to encourage children with autism to recognize and label emotions using a humanoid robot. *23rd IEEE International Symposium on Robot and Human Interactive Communication, IEEE RO-MAN 2014*, 820-825.

[T50] Malik, N.A. Shamsuddin, S. Yussof, H. Miskam, M.A. Hamid, A.C. (2013). Feasibility of using a humanoid robot to elicit communicational response in children with mild autism. *5th International Conference on Mechatronics, ICOM 2013*, 53(1).

[T51] Anamaria, P.C. Ramona, S. Sebastian, P. Jelle, S. Alina, R. Daniel, D. Johan, V. Dirk, L. Bram, V. (2013). Can the social robot probo help children with autism to identify situation-based emotions? A series of single case experiments. *International Journal of Humanoid Robotics*, 10(3).

[T53] Silva, S. Soares, F. Costa, S. Pereira, A.P. Moreira, F. (2012). Development of skills in children with ASD using a robotic platform. *2012 IEEE 2nd Portuguese Meeting in Bioengineering, ENBENG 2012*.
[T54] Scassellati, B. Admoni, H. Matarić, M. (2012). Robots for use in autism research. *ANNUAL REVIEW OF BIOMEDICAL ENGINEERING*, 14, 275-294.
[T55] Vanderborght, B. Simut, R. Saldien, J. Pop, C. Rusu, A.S. Pintea, S. Lefeber, D. David, D.O. (2012). Using the social robot Probo as a social story telling agent for children with ASD. *Interaction Studies*, 13(3), 348-372.
[T56] Giannopulu, I. Pradel, G. (2012). From child-robot interaction to child-robot-therapist interaction: A case study in autism. *Applied Bionics and Biomechanics*, 9(2), 173-179.
[T58] Mazzei, D. Billeci, L. Armato, A. Lazzeri, N. Cisternino, A. Pioggia, G. Igliozzi, R.' Muratori, F. Ahluwalia, A. De Rossi, D. (2010). The FACE of autism. *19th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2010*, 791-796.
[T63] Liu, C. Conn, K. Sarkar, N. Stone, W. (2008). Online affect detection and robot behavior adaptation for intervention of children with autism. *IEEE Transactions on Robotics*, 24(4), 883-896.
[T68] Kozima, H. Nakagawa, C. Yasuda, Y. (2007). Children-robot interaction: a pilot study in autism therapy. *FROM ACTION TO COGNITION*, 164, 385-400.
[T72] Scassellati, B., Boccanfuso, L., Huang, C. M., Mademtzi, M., Qin, M., Salomons, N., ... & Shic, F. (2018). Improving social skills in children with ASD using a long-term, in-home social robot. Science Robotics, 3(21), eaat7544.
[T73] Aziz, A. A., Mokhsin, M., Moganan, F. F. M., Ismail, A., Sakamat, N., Zainol, A. S., & Lokman, A. M. (2018). Humanoid-robot as teaching mediator: Research model in demonstrating the autistic children learning motivation based on the emotional responses. *Advanced Science Letters*, *24*(4), 2296-2300.
[T77] Boccanfuso, L., Barney, E., Foster, C., Ahn, Y. A., Chawarska, K., Scassellati, B., & Shic, F. (2016). Emotional robot to examine different play patterns and affective responses of children with and without ASD. In *11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 19-26). IEEE.
[S1] Park, C. H., Pai, N., Bakthavatchalam, J., Li, Y., Jeon, M., & Howard, A. (2015). Robotic framework for music-based emotional and social engagement with children with Autism. In *AAAI Workshop: Artificial Intelligence Applied to Assistive Technologies and Smart Environments* (pp. 39-40). AAAI.
[S2] Dautenhahn, K., & Werry, I. (2002). A quantitative technique for analysing robot-human interactions. In *IEEE/RSJ International Conference on Intelligent Robots and Systems* (Vol. 2, pp. 1132-1138). IEEE.
[S3] Shamsuddin, S., Yussof, H., Ismail, L. I., Mohamed, S., Hanapiah, F. A., & Zahari, N. I. (2012). Humanoid robot NAO interacting with autistic children of moderately impaired intelligence to augment communication skills. *Procedia Engineering*, *41*, 1533-1538.
[S4] Pour, A. G., Taheri, A., Alemi, M., & Meghdari, A. (2018). Human–robot facial expression reciprocal interaction platform: case studies on children with autism. *International Journal of Social Robotics*, *10*(2), 179-198.
[S6] Billing, E., Belpaeme, T., Cai, H., Cao, H. L., Ciocan, A., Costescu, C., ... & Ziemke, T. (2020). The DREAM Dataset: Supporting a data-driven study of autism spectrum disorder and robot enhanced therapy. *PloS one*, *15*(8), e0236939.

[S7] Costa, S., Soares, F., Pereira, A. P., Santos, C., & Hiolle, A. (2014, October). A pilot study using imitation and storytelling scenarios as activities for labelling emotions by children with autism using a humanoid robot. In *4th International Conference on Development and Learning and on Epigenetic Robotics* (pp. 299-304). IEEE.
[S8] Rudovic, O., Lee, J., Mascarell-Maricic, L., Schuller, B. W., & Picard, R. W. (2017). Measuring engagement in robot-assisted autism therapy: A cross-cultural study. *Frontiers in Robotics and AI*, *4*, 36.

References

[1] Kitchenham B.: Procedures for Performing Systematic Reviews, Technical Report, NICTA Technical Report 0400011T.1, http://www.it.hiof.no/~haraldh/misc/2016-08-22-smat/Kitchenham-Systematic-Review2004.pdf
[2] Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097
[3] A. Landowska, A. Karpus, T. Zawadzka, B. Robins, D. Erol Barkana, H. Kose., T. Zorcec and N. Cummins., "Emotion Recognition in Children with Autism: A Systematic Literature Review" (under preparation)
[4] Ishihara S., Nagamachi M., Schütte S., Eklund J.: Affective meaning: The Kansei engineering approach, Product Experience, 2008, 477-496, https://doi.org/10.1016/B978-008045089-6.50023-X
[5] Kansei engineering linking emotions and features https://aspzbwii.wordpress.com/2009/10/06/kansei-engineering-linking-emotions-and-product-features-tlumaczenie/