Project No. 611391

DREAM

Development of Robot-enhanced Therapy for Children with Autism Spectrum Disorders

Agreement Type: Collaborative Project
Agreement Number: 611391

D1.1 Intervention Definition

Due Date: **01/06/2014**
Submission date: **23/06/2014**

Start date of project: **01/04/2014** Duration: **54 months**

Organisation name of lead contractor for this deliverable: **Universitatea Babes Bolyai**

Responsible Person: **Daniel David**

Revision: **7.2**

Project co-funded by the European Commission within the Seventh Framework Programme		
Dissemination Level		
PU	Public	**PU**
PP	Restricted to other programme participants (including the Commission Service)	
RE	Restricted to a group specified by the consortium (including the Commission Service)	
CO	Confidential, only for members of the consortium (including the Commission Service)	

Contents

I Executive Summary

Deliverable D1.1 in Development of Robot-enhanced Therapy (RET) for Children with Autism Spectrum Disorders project is concerned with the definition of the interventions to be carried out between a child with ASD, a therapist, and the robot. It will form part of the Reference Manual of Clinical Requirements and it describes the detailed scenarios for the three types of intervention: (a) Joint attention, (b) imitation, and (c) turn-taking that are key parts of robot-enhanced therapy (RET).

D1.1 is important because it provides the basis for the robot behaviour specification (D1.2) and the child behaviour specification (D1.3). In turn, D1.2 and D1.3 provide the essential definition of requirements for the work to be done in Work Packages 4, 5, and 6: Sensing and Interpretation, Child Behaviour Analysis, and Robot Behaviour, respectively.

II Principal Contributors

The main authors of this deliverable are as follows (in alphabetical order).

Cristina Costescu, Babes-Bolyai University
Daniel David, Babes-Bolyai University
Anca Dobrean, Babes-Bolyai University
Honghai Liu, University of Portsmouth
Silviu Matu, Babes-Bolyai University
Aurora Szentagotai, Babes-Bolyai University
Serge Thill, University of Skövde
David Vernon, University of Skövde

III Revision History

Version 1.0 (DV 04-06-2014)
First draft.

Version 2.0. (CC 04-06-2014)
Second draft

Version 3.0. (DD, AS, AD, SM 06-06-2014)
Third draft (integrating the first and the second draft).

Version 4.0 (DD, CC, AS, AD, SM 09-06-2014)
Fourth draft.

Version 5.0 (DD 09-06-2014)
Final draft.

Version 6.0 (DV 17-06-2014)
Added a specification of action the robot is to take when the child exhibits unexpected behaviour (Section 3.9). Added an explanation of the camera configuration (Section 4).

Version 7.0 (DD 19-06-2014)
Removed level 3 from Section 3.8: Turn-taking Intervention, and expanded Level 2 to deal with rational statements and adaptive strategy.

Version 7.1 (DV 23-06-2014)
Fix some minor typos.

Version 7.2 (DV 01-07-2014)
Added note to Section 4, Therapy Environment, Subsection System Design, to say that the final sensor configuration will be based on the application requirements.
Changed the reference to a "workbench" to "therapy table", throughout. The robot sits either at or on the therapy table, depending on the situation.
Section 2 Overview no longer states that Section 4 identifies an inventory of supporting material and objects to be used in the interventions.

IV Intervention Definition

1 Introduction: Theoretical Framework

Fundamentals

Autism Spectrum Disorder (ASD) is one of the most common childhood developmental disorders (Fombonne, 2009). ASD is characterized by restricted patterns of behavior and interests and qualitative impairments in communication and social interaction (American Psychiatric Association, 2013). These are known collectively as the core symptoms for ASD (American Psychiatric Association, 2013).

These core symptoms emerge early and persist in development even though their precise manifestation changes over the course of development. Although during adolescence and adulthood the core symptoms of ASD seem to decrease over time, ASD is seen as having chronic disability with a poor prognosis pertaining development. Most individuals with ASD require professional care throughout their lives (Howlin, Goode, Hutton, & Rutter, 2004; Mordre, Groholt, Knudsen et al., 2012).

Currently, no clear biological marker exists for ASD. The diagnostic criteria for Autism Spectrum Disorders included in the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5) (American Psychiatric Association, 2013), refer to Autism Spectrum Disorder as a single diagnosis category. More specific, the criteria refer to deficits into two categories: (1) Social Communication domain (e.g., persistent deficits in social communication and social interaction across contexts, not accounted for by general developmental delays) and (2) Restricted, repetitive patterns of behavior, interests, or activities. In terms of assessment and diagnostic process, ASD children are identified based on the behavioral phenotype, *Autism Diagnostic Interview-Revised* (ADI-R) (Lord, Rutter, & Le Couteur, 1994) and *The Autism Diagnostic Observation Schedule* (ADOS) (Lord, Rutter, Goode et. al., 1989) being used by researchers and academic centers as golden standards.

Moreover, causal mechanisms are not well understood and/or integrated into a rigorous etiopatogenetic theory, although several hypotheses have been advanced. For example, Courchesne and colleagues (2001, 2003, 2007) conceptualize autism as involving two phases of early brain growth pathology: early brain overgrowth at the beginning of life and slowing or arrest of growth during early childhood. According to the *Empathizing-Systemizing (E-S) Theory* of psychological sex differences, proposed by Baron-Cohen, (e.g., Baron-Cohen, 2009; Baron-Cohen, Knickmeyer, & Belmonte, 2005), human males have stronger systemizing tendencies (i.e., analyzing a system in terms of the rules that govern it, in order to predict its behavior) compared to females, who exhibit stronger empathizing tendencies (i.e., the drive to identify another's mental states and to respond to them appropriately). These differences are brain-structure and function based. An extension of this theory, the *Extreme Male Brain Theory (*EMB), proposes that individuals with ADS are characterized by impairments in empathizing alongside intact or even superior systemizing (Baron-Cohen, Wheelwright, Hill, et al., 2001). These impairments could be explained by prenatal exposure to atypically high levels of androgens (e.g., testosterone). Basically, from a psychological point of view, most of the research and interventions were focused either on more basic mechanisms (e.g., imitation), or various basic cognitive-behavioral skills (e.g., joint attention and turn taking), both related to the core symptoms (e.g., communication and social interaction) and disability/impairments (e.g., lacking social skills like crossing the street etc.). Indeed, severity of autism is correlated with impaired imitation skills, joint attention, and turn taking (Rogers, Hepburn, Stackhouse, & Wehner, 2003). As such, children with autism fail to imitate and to have joint attention episodes from an early age and this lack of imitation and joint attention is a salient diagnostic marker for the disorder (Lord, Risi, Lambrecht et al., 2000). Thus, although the mechanisms are not yet clearly understood, we know that various biological and psychological "causes" [basic mechanisms (e.g., imitation) and basic cognitive-behavioral skills (e.g., joint attention and turn taking)] generate the core symptoms of ASD that are then often related to impairment/disability.

For over 40 years, researchers have explored how to clinically help children with ASD (Dawson & Adams, 1984; Lovaas, 1987; Lovaas, Freitas, Nelson, & Whalen, 1967; Rogers, Bennetto, McEvoy, & Pennington, 1996). For children with ASD, better imitation skills, joint attention, and turn taking appear to be related to improved language performance (Stone, Ousley, & Littleford, 1997), play skills (Smith, Mirenda, & Zaidman-Zait, 2007), and social skills (Carpenter, Pennington, & Rogers, 2002; Ingersoll, 2011). These skills are taught usually in discrete, analogue settings in adult-child exchanges (Cardon & Wilcox, 2011; Ingersoll & Schreibman, 2006; Lovaas et al., 1967). Although some skill acquisition occurred in these settings, the development of these skills was extremely limited.

Recently, researchers have found that for children with ASD, imitation and joint attention acquisition improves in settings in which technological tools are involved (Scassellati, Admoni, & Mataric, 2012; Ricks & Colton, 2010; Michaud & Clavet, 2001; Robins, Amirabdollahian, Ji, & Dautenhahn, 2010; Kozima, Nakagawa, & Yasuda, 2005; Vanderborght, Simut, Saldien et al, 2012; Tapus, Peca, Aly et al, 2012).

Different theories try to explain why children with autism prefer to interact with technological tools. One of them, the Theory of Mind (TOM) (Baron-Cohen, 1997) explains that children with autism tend to have difficulties in identifying the mental states of others, i.e. in having a representation of what others may think. More precisely, TOM refers to a full range of mental states (e.g., beliefs, desires, intentions, imagination, emotions) that cause action [for a description of some of the manifestations of this impairment see Baron-Cohen (1997)]. Consequently, it can be very hard for them to understand social human-human interactions and thus, they prefer technological tools in order to live in a predictable world. Moreover, they often lack the capability to generalize (Baron-Cohen, 1997) and, as a consequence, to classify entities. Furthermore, children with autism can distinguish between a human and an object but, their behaviour towards humans may have elements of how they treat objects (Hobson, 2002). Moreover, since human beings are very complex with all their essential expressiveness, children with autism tend to prefer interacting with objects which are simpler and more predictable. Indeed, social situations contain an incredible amount of information, very difficult for the child to systemize and therefore, to understand it. This could be partly explained by a theory focusing on the empathizing–systemizing of Baron-Cohen (Baron-Cohen, 2009).

Taking into account that ASD patients tend to learn more from the interaction with technology rather than from the interaction with the human beings, robots might have the potential to be used in ASD therapies as a mediators between human models and ASD patients (see David, Matu, & David, 2014). In the Robo-Mediator approach, the robot is used as a mean for delivering the treatment because it enables faster and better gains from the therapeutic intervention as compared to the classical condition, in which there is only direct interaction between therapist and patient. The robot acts as a necessary component in the process and without it the treatment would attain poorer results, or it would take longer to attain same results. In this case, ASD patients might develop faster, or even to a greater performance level, the relevant skills (e.g., imitation, joint attention, turn taking) in a psychological intervention mediated by the robot because they find it easier to interact with such agents than with human agents. The Robot-Mediator approach is different from other integrations of robotics in psychotherapy. In the case of the Robot-Therapist approach, the robot acts by itself as the therapist and completely replaces the human agent. In the case of the Robo-Assistant, the robot acts as a facilitator of the process, but is not a crucial or necessary component for treatment success, and could be easily replaced by other agents (e.g., animal agents, peers). All these means of integrating robotic agents into psychological interventions are not new forms of treatment, but rather new ways of delivering the same treatment. They all make use of the same theory on the psychological problems and the same treatment principles, but use different roles for the robotic agent to deliver the intervention (for more details see David et al., 2014).

The choice for a robot-mediated approach to psychological intervention for ASD children is justified by several advantages: 1) Children with ASD are more responsive to feedback, even social feedback, when administered via technology rather than a human (Ozonoff, 1995); 2) The anthropomorphic embodiment of the robot offers human like social cues, while keeping at the same time object-like simplicity; 3) Robots can be programmed to gradually increase the complexity of the tasks, by solely presenting relevant information; moreover, information can be repeated in the same format, without trainer fatigue; 4) Robots are predictable and, therefore, controllable, enable errors to be made safely and give possibilities to train a wide range of social and communication behaviours to prepare for real life exposition;

The clinical application of the DREAM [Development of Robot-enhanced Therapy (RET) for Children with Autism Spectrum Disorders] project aims to investigate how application to investigate how children with ASD behave and how they perform when interacting with the Nao robot, compared to a human partner in an imitation task, joint attention task, and turn taking. Thus, the robot is created and tested as a tool to develop imitation, joint attention, and turn taking in ASD patient, with the final aim of using these developments for a better real life social interaction of the ASD children.

We have defined the measured variables that in the child-robot interaction as follow (for more details regarding the way in which these behaviours will be measures see the below sections 2-5):

Imitation task

– Imitation of the movements made by the robot, either with the objects or without (in terms of frequency);
– Accuracy of the imitation of the movements made by the robot with the objects or without;
– Initiations of motor actions that the child performs, which are triggered by the robot or by a soft physical prompt (i.e., softly touching the elbows of the child for one second);
– Imitation of the gestures made by the robot, gestures that refer to a specific emotion (one out of four: anger, happiness, sadness and fear) and include hand movements and head movements (in terms of frequency);
- Imitation of the sounds made by the robot, sounds that accompany a specific emotion (one out of four: anger, happiness, sadness and fear);
- Showing the correct facial expression that accompanies the emotion reproduced by the robot (in terms of frequency).

Joint attention

- Gaze alternating consisting of the child independently alternating his gaze (i.e., looking at the target object, at the robot, and back at the target object) within 4 s of the presentation of the discriminative stimulus (SD) – which in our case will be a picture;
- Gaze alternating and pointing consisted of the child independently, within 4 s of the SD, alternating his gaze and pointing (i.e., extending his arm and index finger in the direction of the object/event), either simultaneously with, or immediately following, gaze alternating;
- Gaze alternating, pointing, and verbalizing consisted of the child independently, within 4 s of the SD alternating his gaze, pointing, and verbalizing (either simultaneously with, or immediately following, gaze alternating and pointing);
– Showing the correct gestures that refer to a specific emotion (one out of four: anger, happiness, sadness and fear) and include hand movements and head movements (in terms of frequency);
- Showing the correct facial expression that accompanies the emotion reproduced by the robot in relation with the picture content (in terms of frequency).

Turn taking

- Exchanging information with the robot: contingent utterances - verbal utterances (one word or a couple of words) that are in context, congruous with the interaction partner (e.g., yes-no

responses, responses to the question) (measured in frequency – the number of contingent utterances said by the child during the task);
- Rational and irrational beliefs – analyses of the discourse that the child has during the interaction with the robot;
- Expressions of anger and sadness during the robot task - these emotions (anger, happiness, sadness and fear) will be coded on the basis of facial, vocal, or postural cues;
- Adaptive and maladaptive behaviours as a solution for the social scenarios presented by the robot.

Besides these primary outcomes (i.e., specific response measurements), we also have some secondary outcomes that are relevant for every session with the robot, regardless the specific task:

Stereotypical behaviours: a repetitive or ritualistic movement, posture, or utterance (measured in frequency – the number of stereotype behaviours performed by the child during the task).

Positive emotions: the child laughed or smiled while interacting with the robot (measured in frequency - the number of smiles or laughs performed by the child during the task).

Contingent utterances: verbal utterances (one word or a couple of words) that are in context, congruous with the interaction with the partner (e.g. yes-no responses, responses to the question) (measured in frequency – the number of contingent utterances said by the child during the task).

Verbal initiations: verbal utterances (one word or a couple of words) that are in context, congruous with the interaction with the robot and adds a new information, including expansion, adding to the content of the robot utterance or introducing new related topics (e.g. ask some questions, makes references to their own personal experience (measured in frequency – the number of verbal initiations made by the child during the task).

Eye contact: looking at the upper region (not necessary at the eyes) of the robot for more than 3 seconds (measured in duration – the number of seconds in which the child made eye-contact with th robot).

Engagement in the task: (see the below in Table 1 the rating system that we used before Pop, Pintea, Vanderborght & David, 2014)

Table 1. The rating system for engagement task (after Pop, Pintea, Vanderborght, & David, 2014).

Rating	***Meaning***	***Description***
0	*Intense noncompliance*	*The child walked away from the place in which the robot/adult interaction took place*
1	*Noncompliance*	*The child refused to comply with the experimenter's request to play with the robot/adult*
2	*Neutral*	*The child complied with instructions to help the robot/adult after several prompts from the experimenter.*
3	*Slight interest*	*The child required two or three prompts from the experimenter before responding to the robot/adult.*
4	*Engagement*	*The child complied immediately following the experimenter's request to help the robot/adult.*
5	*Intense engagement*	*The child spontaneously engaged with the robot/adult.*

2 Overview

Deliverable D1.1 is concerned with the definition of the interventions to be carried out between a child with ASD, a therapist, and the robot. It will form part of the Reference Manual of Clinical Requirements and it describes the detailed scenarios for the three types of intervention (a) joint attention, (b) imitation, and (c) turn-taking that are key parts of robot-enhanced therapy (RET).

D1.1 is important because it provides the basis for the robot behaviour specification (D1.2) and the child behaviour specification (D1.3). In turn, D1.2 and D1.2 provide the essential definition of requirements for the work to be done in Work Packages 4, 5, and 6: Sensing and Interpretation, Child Behaviour Analysis, and Robot Behaviour, respectively.

The project Description of Work states that this deliverable will include a walk-through of all forms of the three types of intervention. Originally, it was envisaged that there would be 14 exercises, based on different levels in each intervention and two different types of imitation intervention, each with its own levels; see Figure 1.

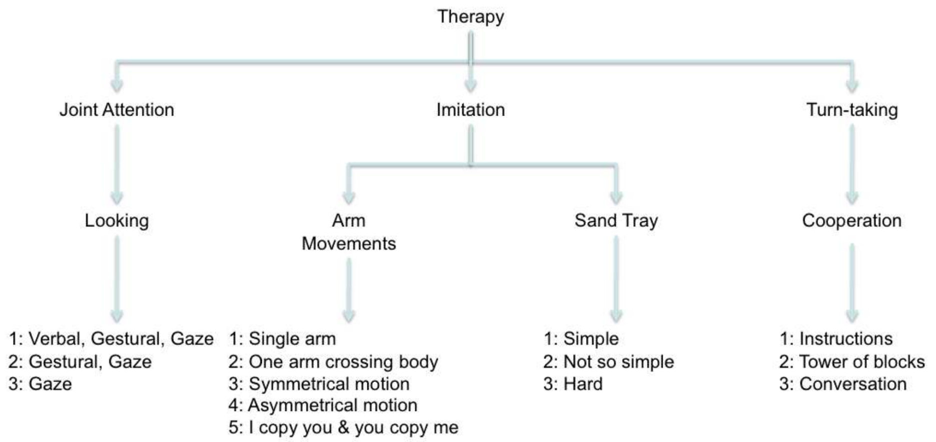

Figure 1: The fourteen exercises comprising the three types of intervention.

The Description of Work also states the each exercise would be broken down into a time sequence of elementary robot actions and that, for each action, the deliverable would specify:

1. The set of triggers for the action (e.g. input from therapist, child, or environment)
2. The sensory cues that characterize each trigger;
3. The exact sequence of movements, expressions, or vocal output that constitute the action and their associated sensory cues;
4. The goal of the action, i.e. the expected change in the environment, the response of the child of the therapist;
5. The sensory cues that characterize the goal of the action; in each case, there may be multiple triggers and responses.

The purpose of this decomposition was to identify explicitly all the sensory-motor requirements of the robot. However, it was subsequently decided to define the intervention in a slightly different, but fully compatible, manner. Specifically, the interventions are now described as a sequence of actions, each action comprising a number of component movements and sensory cues linked to a particular sensory-motor process. The motor aspect of these processes provides the basis for the robot behaviour specification in D1.2 and hence the robot behaviour in processes that have to be developed in Work Package 6. The sensory aspect provides the basis for the child behaviour specification in D1.3 and hence the sensing and interpretation processes that have to be developed in Work Package 4, as well as the child behaviour analysis in Work Package 5.

The fourteen exercises anticipated in the Description of Work have now been consolidated into an alternative form of intervention definition, comprising nine tasks as follows.

1. A general set of actions used to start all tasks.
2. A joint attention diagnosis task with three steps.
3. A joint attention intervention task with two phases: robot-initiated and child-initiated.
4. An imitation diagnosis task *with* objects and two phases:
    a. functional imitation (four movements & four objects);
    b. symbolic imitation (four movements & one object).
5. An imitation diagnosis task *without* objects (four movements).
6. An imitation intervention task *without* objects (four emotions).
7. A turn-taking diagnosis task.
8. A turn-taking intervention task with two levels.
9. A set of actions the robot is to take when the child exhibits unexpected behaviour.

Regarding the ninth task, this situation will arise when, for example, the child becomes bored and doesn't interact in the way that the intervention or diagnosis anticipates, i.e. when the intervention or the diagnosis doesn't go according to plan. This includes the situation where the robot is waiting for the child's interaction but the child doesn't do anything. These situations are handled by providing an implicit fail-safe action that is invoked if the expected condition for robot-child interaction isn't met by some pre-specified time interval.

Section 3 provides the detailed decomposition of these nine sets of tasks.

Section 4 describes the environment in which the child, robot, and therapist will work.

Finally, Section 5 provides an outline design of the therapy table to support the high-resolution video cameras, the Kinect RGB-D cameras, and the robot itself.

3 Definition of the Interventions

This section comprises eight sub-sections, each one setting out a detailed decomposition of the eight forms of diagnosis and intervention described in the previous section. These decompositions are relatively straightforward but do contain a lot of information. To make it easier to read, the parts of the decomposition – actions, movements and sensory cues, sensory-motor processes, and comments – are colour-coded.

The actions (in green) define the intervention and diagnosis tasks in a relatively abstract and intuitive manner.

The component movements and sensory cues (in black) develop this to make explicit all of the constituents of each action.

The component movements and sensory cues always refer to the robot's perspective, i.e. they define what the robot does and what the robot sees and hears.

The sensory-motor processes (in red) provide the essential input for the definition of robot behaviour specification (D1.2) and the child behaviour specification (D1.3).

The comments (in blue) add simple explanations of what is happening in the task at that point.

3.1 Actions at the start of all RET tasks

The robot sits on the table, actively waiting to engage

REPEAT	
The robot periodically makes natural movements	
The robot listens for a loud noise	
The robot hears a loud noise	Detection of high amplitude sounds
The robot locates the sound	Sound localization in horizontal plane
The robot moves its head to look at the sound	Move head to centre gaze on the sound
The robot looks for the therapist	
Detects a face in its field of vision	Face detection
If it is the therapist's face, the robot	Face recognition
Determines the position of the therapist	Face localization
Moves its head to gaze at the therapist	Move head to centre gaze on the therapist
Adjusts its body posture to face therapist	Move torso to face therapist and adjust gaze
Maintains its gaze for a short period of time	Active face tracking
The robot looks for the child	
Detects a face in its field of vision	Face detection
If it is the child's face, the robot	Face recognition
Determines the position of the child	Face localization
Move its head to gaze at the child	Move head to centre gaze on the child
Adjust its body posture to face the child	Move torso to face child and adjust gaze
Maintains its gaze for a short period of time	Active face tracking
UNTIL the child is sitting in front of the robot	Child body pose recognition
The robot engages with the child	
The robot says "Hello! I am glad you are here to play with me"	Speech synthesis: say "Hello! I am glad you are here to play with me"
The robot announces that the game is about to begin	
IF the child is engaged and paying attention	Detect mutual gaze (Note 1) Recognize facial expression (Note 2)
The robot says "Today, we will play together"	Speech synthesis: say "Today, we will play together"

Notes

1 When detecting mutual gaze, the robot only has to determine whether or not the child looks at the robot's head, but not necessarily at the robot's eyes. The gaze has to be held for a minimum period, e.g. 3 seconds.

2 When detecting that the child is paying attention based on the child's facial expression, it is sufficient to detect an expression of interest: focussed gaze and neutral expression

3.2 Joint Attention Diagnosis ADOS

Pictures are placed to the left and the right of the robot, facing the child

There are three steps:

Step 1: gaze only
Step 2: gaze and pointing
Step 3: gaze, pointing, and vocal instruction

REPEAT	
Check to see that the child is looking at the robot	
REPEAT	
Look for a face	Face detection
UNTIL the child's face is detected	Face recognition
Determine the location of the child	Face localization
Move head to look at the child	Move head to centre gaze on the child
Adjust body posture to face the child	Move torso to face child and adjust gaze
REPEAT	
Look at the child	Face localization
UNTIL the child is looking at the robot	Detect mutual gaze
Look at a picture	
The robot looks at the picture to the left or right	Object (picture) detection
	Object (picture) localization
	Move head to centre gaze on the picture
The robot stares at the picture for a specified time	
Point at a picture	
IF at Step 2 (i.e. gaze & pointing)	
The robot points at the picture	Move arm to point at the picture
Say "Look"	
IF at Step 3 (i.e. gaze, pointing, and vocal instruction)	
The robot points at the picture	Move arm to point at the picture
The robot says "Look at the <object> in the picture"	Object (picture) recognition
Look back at child	
REPEAT	
Look for a face	Face detection
UNTIL the child's face is detected	Face recognition
Determine the location of the child	Face localization
Move head to look at the child	Move head to centre gaze on the child
Adjust body posture to face the child	Move torso to face child and adjust gaze
Check to see if the child is looking at the picture	
Determine the gaze direction of the child	Compute child's head gaze
	Compute child's eye gaze (ideally)
Wait for the child to return its gaze to the robot	
REPEAT	
Look at the child	Face localization
UNTIL the child looks at the robot	Detect mutual gaze
Wait 10 seconds	
If the child did not gaze at the picture, make a second attempt	
IF second attempt failed in Step 1 (i.e. child does not look at the correct picture)	
Proceed to Step 2: gaze & pointing	
IF second attempt failed in Step 2 (i.e. child does not look at the correct picture)	
Proceed to Step 3: gaze, pointing, and speech	
UNTIL the child gazes at the correct picture	

3.3 Joint Attention Intervention

Pictures are placed to the left and the right of the robot, facing the child
There are two phases: in phase 1 the robot initiates, in phase 2 the child initiates

REPEAT

Phase 1: the robot initiates the interaction
The robot looks at a picture: the child follows the robot's gaze and enacts the emotion in the picture

Check to see that the child is looking at the robot	
REPEAT	
Look for a face	Face detection
UNTIL the child's face is detected	Face recognition
Determine the position of the child	Face localization
Move head to look at the child	Move head to centre gaze on the child
Adjust body posture to face the child	Move torso to face child and adjust gaze
REPEAT	
Look at the child	Face localization
UNTIL the child is looking at the robot	Detect mutual gaze
Shift attention to a picture	
The robot looks at the picture to the left or right	Object (picture) detection
The robot recognizes the emotion in the picture	Object (picture) recognition
The robot holds its gaze on the picture for a pre-specified time	
Show an emotion based on the content of the picture	
IF picture is angry the robot strikes an angry pose	
IF picture is happy the robot strikes an happy pose	
IF picture is sad the robot strikes an sad pose	
IF picture is scary the robot strikes an fearful pose	
Look back at the child	
REPEAT	
Looks for a face	Face detection
UNTIL the child's face is detected	Face recognition
Determine the position of the child	Face localization
Move head to gaze at the child	Move head to centre gaze on the child
Adjust body posture to face the child	Move torso to face child and adjust gaze
Check to see if the child is looking at the picture	
Determine the gaze direction of the child	Compute child's head gaze
	Compute child's eye gaze (ideally)
Wait for the child to return its gaze to the robot	
REPEAT	
Determine the gaze direction of the child	Face localization
UNTIL the child looks at the robot	Detect mutual gaze
Watch the child enact the emotion and check if it is correct	
Track the child’s hands and head	Head and hand tracking
Classify the child's body pose	Child body pose recognition
Classify the child's facial expression	Facial expression recognition

Phase 2: the child initiates the interaction
The child looks at a picture: the robot follows the child's gaze and enacts the emotion in the picture

Look at the child	
REPEAT	
Detect a face	Face detection
UNTIL the child's face is detected	Face recognition
Determine the position of the child	Face localization
Move head to gaze at the child	Move head to centre gaze on the child
Adjust body posture to face the child	
The child looks at a picture and the robot follows its gaze	
REPEAT	
Look at the child	Face localization
UNTIL the child is looking at the picture	Compute child's head gaze Compute child's eye gaze (ideally)
The robot looks where the child is looking	Determine intersection of gaze and table Move head to centre gaze on this area Search this area for a picture Object (picture) detection Object (picture) localization
The robot recognizes the emotion in the picture	Object (picture) recognition
The robot displays an emotion based on the content of the picture	
IF picture is angry the robot strikes an angry pose	Move to a pre-determined pose
IF picture is happy the robot strikes an happy pose	
IF picture is sad the robot strikes an sad pose	
IF picture is scary the robot strikes an fearful pose	

UNTIL the child has looked at both pictures

3.4 Imitation Diagnosis with Objects

Phase 1: Functional Imitation

REPEAT

Select current movement:
(1) sliding a car on a table
(2) drinking from a cup
(3) waving an airplane in the air
(4) smelling a flower

The robot looks at the table, sees an object, and recognizes it	
The robot looks at the table	Move head to centre gaze on the table
The robot looks at the object	Object detection
	Object localization
	Move head to centre gaze on the object
The robot recognizes the object	Object recognition
The robot picks up the object	
Determine the grip point	Grip point localization
IF the grip point is on top of the object	
Reach to the top of the object	Move hand above grip location
Move hand down in contact with the object	Move hand down
Activate gripper (e.g. electromagnet)	Grip object (activate electromagnet)
IF the grip point is at the side of the object	
Reach to the right side of the object (right-handed robot)	Move hand right of grip location
Move hand left in contact with the object	Move hand left
Activate gripper (e.g. electromagnet)	Grip object (activate electromagnet)
Lift the object	Move hand up

The robot moves the object around, demonstrating the action to be imitated

Car: slide left and right on table and say broomm!
Aeroplane: movel left and right in the air and say zoomm!
Flower: lift to touch face
Cup: lift to touch face

IF object is a car	
Lower hand & object so that object is in contact with table	Move hand down
Move object left & right and make an engine sound	Move hand left; Say "Broomm"
	Move hand right Say "Broomm"
	Move hand left; Say "Broomm"
	Move hand right Say "Broomm"
IF object is an aeroplane	
Move object left & right and make an engine sound	Move hand left; Say "Zoomm"
	Move hand right Say "Zooomm"
	Move hand left; Say "Zoomm"
	Move hand right Say "Zooomm"
IF object is a flower or a cup	
Lift object to touch robot's face	Move up to face
	Wait one second
Place object back at the original location	Move hand above original grip location
	Move hand down
	Release object (deactivate electromagnet)

The child should now take the object from the table

The robot looks to see if the child takes the object
REPEAT
Detect a face
UNTIL the child's face is detected
Determine the position of the child
Move head to gaze at the child
Detect the child's hand
Determine the position of the hand
REPEAT
Look at the hand
UNTIL the hand is close to object

REPEAT
Look at the object
UNTIL the object is removed from the table

Face detection
Face recognition
Face localization
Move head to centre gaze on the child
Hand detection
Hand localization

Hand tracking
Object detection
Object localization
Hand-object spatial relationship
Object-table spatial relationship

Object tracking
Hand-object spatial relationship

The child should now imitate the actions

The robot looks to see if the child moves the object correctly
REPEAT
Look at the object
Compare object movement with expected pattern
UNTIL the child puts it back on the table

Object tracking
Object tracking
Trajectory classification
Hand tracking
Hand-object spatial relationship

Wait 10 seconds and repeat the exercise with the same object.

Wait 10 seconds

UNTIL all four movements have been completed

Phase 2: Symbolic Imitation

REPEAT

Select current movement (any object)
(1) sliding
(2) drinking
(3) waving
(4) smelling

Perform exactly the same actions as for the functional imitation

UNTIL all four movements have been completed

3.5 Imitation Diagnosis without Objects

REPEAT

Select current movement:
(1) Covers eyes
(2) Touches head with two hands
(3) Airplane arms
(4) Waving with one hand

Check to see that the child is looking at the robot

REPEAT	
Look for a face	Face detection
UNTIL the child's face is detected	Face recognition
Determine the position of the child	Face localization
Move head to look at the child	Move head to centre gaze on the child
Adjust body posture to face the child	Move torso to face child and adjust gaze
REPEAT	
Look at the child	Face localization
UNTIL the child is looking at the robot	Detect mutual gaze

The robot makes the appropriate movement

The robot covers its eyes, or	Move to a pre-determined pose
The robot touches it head with both hands, or	
The robot stretches out its arms, or	
The robot waves with one hand	

The child executes the movement

The robot looks to see whether the child does the same movement

Detect the child's hand	Hand detection
Determine the position of the hand	Hand localization
Move head to look at the hand	Move head to centre gaze on the hand
REPEAT	
Look at the hand	Hand tracking
UNTIL	
(1) the child's hand covers to the child's eyes, or	Hand-object (eye) occlusion detection
(2) the child's hand is close to the child's head or	Hand-object (head) spatial relationship
(3) the child's hand is extended horizontally, or	Hand-object (body) spatial relationship
(4) the child's hand waves back and forth	Hand-object (body) alignment detection
(5) a fixed period of time passes	Trajectory classification

If a fixed period of time has passed, then the child has not imitated the action so repeat the current movement just one more time

UNTIL all four movements have been completed

3.6 Imitation Intervention without Objects

REPEAT

Select current emotion:
(1) Happy
(2) Sad
(3) Anger
(4) Fear

Phase 1: the robot initiates and shows the child what emotion to enact

Check to see that the child is looking at the robot

REPEAT	
Look for a face	Face detection
UNTIL the child's face is detected	Face recognition
Determine the position of the child	Face localization
Move head to look at the child	Move head to centre gaze on the child
Adjust body posture to face the child	Move torso to face child and adjust gaze
REPEAT	
Look at the child	Face localization
UNTIL the child is looking at the robot	Detect mutual gaze

The robot displays the required emotion

IF robot is angry the robot strikes an angry pose	Move to a pre-determined pose
IF robot is happy the robot strikes an happy pose	Make a pre-determined sound
IF robot is sad the robot strikes an sad pose	
IF robot is afraid the robot strikes an fearful pose	

The child imitates the movement

Watch the child enact the emotion and check if it is correct

REPEAT	
Track the child's hands and head	Head and hand tracking
Classify the child's body pose	Child body pose recognition
Classify the child's facial expression	Facial expression recognition
UNTIL	
(1) the child makes an attempt to imitate, or	
(2) a fixed period of time passes	

If the child makes no attempt, repeat the current movement just one more time

Assess the child's attempt

IF attempt is good	
Provide very positive feedback	Say "Great job. Well done!"
IF attempt is okay	
Provide positive feedback	Say "Good try!"
IF attempt is poor	
Provide encouraging feedback	Say "Not bad!, Let's try again!"

Phase 2: the child initiates and tells the robot what emotion to enact

Look at the child

REPEAT	
Detect a face	Face detection
UNTIL the child's face is detected	Face recognition
Determine the position of the child	Face localization
Move head to gaze at the child	Move head to centre gaze on the child
Adjust body posture to face the child	

The child tells the robot to be sad, happy, afraid, or angry

The robot listens and understands what the child says

REPEAT	
Listen to the child	Speech recognition
UNTIL the child say " Be sad", "Be happy", "Be afraid, or "Be angry"	
The robot recognizes the emotion	

The robot displays the right emotion
IF the child said “Be happy” the robot strikes a happy pose
IF the child said “Be sad” the robot strikes a sad pose
IF the child said “Be afraid” the robot strikes a fearful pose
IF the child said “Be angry” the robot strikes an angry pose

Move to a pre-determined pose
Make a pre-determined sound

UNTIL all four emotions have been completed

3.7 Turn-taking diagnosis

A family scenario will be illustrated on the sand-tray, including characters and objects

The robot and the child have to use these characters and objects to illustrate a story and some appropriate actions

REPEAT	
Check to see that the child is looking at the sand-tray	
REPEAT	
Look for a face	Face detection
UNTIL the child's face is detected	Face recognition
Determine the position of the child	Face localization
Move head to look at the child	Move head to centre gaze on the child
Adjust body posture to face the child	Move torso to face child and adjust gaze
REPEAT	
Determine the gaze direction of the child	Compute child's head gaze
UNTIL the child is looking at the sand-tray	Compute child's eye gaze (ideally)
Look at the sand-tray	
The robot looks at the sand-tray in front of it	Object (sand-tray) detection Object (sand-tray) localization Move head to centre gaze on the sand-tray
Find an object or a character in the sand-tray picture	
The robot recognizes an object or character	Object (picture) detection Object (picture) localization Object (picture) recognition Move head to centre gaze on the sand-tray
Move object or character in the sand-tray	
The robot touches the object or character	Move hand to touch at the object (picture)
Find a suitable location to put the object or character	Object (destination) detection Object (destination) localization Object (destination) recognition
Move the object or character	Move hand to touch position
Say "Look; now it's your turn"	Speech synthesis: say "Look; now it's your turn "
The robot looks to see whether the child moves an object or character	
Detect the child's hand	Hand detection
Determine the position of the hand	Hand localization
Move head to look at the hand	Move head to centre gaze on the hand
REPEAT	
Look at the hand	Hand tracking
UNTIL	
(1) the child's hand moves an object in the sand-tray, or	Hand-object (picture) spatial relationship
(2) a fixed period of time passes	Hand-object (picture) occlusion detection
If a fixed period of time has passed, then the child has not moved an object so say something encouraging	
Say "It's still your turn … have a go"	Speech synthesis: say "It's still your turn … have a go"
Wait a fixed period of time	Wait
UNTIL the therapist says "Let's stop the game now"	Voice recognition (therapist) Speech recognition

3.8 Turn-taking Intervention

Level 1: Sharing social information

REPEAT	
First the robot says something about itself	
Say "I like to play with children"	Speech synthesis: say "I like to play with children"
Say "It's your turn: what do you like to do?"	"It's your turn: what do you like to do?"
The robot waits for the child to reply REPEAT	
Listen to the child	Voice recognition
UNTIL the child says something	
UNTIL a specified number of sentences have been spoken	Voice recognition (therapist) Speech recognition

Level 2: Social knowledge

Watching a social scenario on the sand-tray, the robot observes a situation that can be related to one of four emotions

(1) Happiness
(2) Sadness
(3) Anger
(4) Fear

REPEAT

Phase 1: the robot initiates and chooses an emotion-related picture

Check to see that the child is looking at the sand-tray	
REPEAT	
Look for a face	Face detection
UNTIL the child's face is detected	Face recognition
Determine the position of the child	Face localization
Move head to look at the child	Move head to centre gaze on the child
Adjust body posture to face the child	Move torso to face child and adjust gaze
REPEAT	
Determine the gaze direction of the child	Compute child's head gaze
UNTIL the child is looking at the sand-tray	Compute child's eye gaze (ideally)
Look at the sand-tray	
The robot looks at the sand-tray in front of it	Object (sand-tray) detection Object (sand-tray) localization Move head to centre gaze on the sand-tray
The robot recognizes the scenario depicted in the sand-tray picture	
The robot recognizes the sand-tray picture	Object (picture) recognition
The robot recalls the associated story from memory	
The robot describes the scenario	Say “<scenario description>”

Several pictures denoting different emotions appears on the sand-tray

The robot recognizes the picture	
The robot recognizes the sand-tray picture	Object (picture) recognition
The robot says that’s happy, sad, afraid, or angry	Say “That’s <emotion>”

Phase 2: the child has to choose an emotion-related picture and say what it is

Several pictures denoting different emotions appears on the sand-tray

The robot listens and understands what the child says	
REPEAT	
Listen to the child	Speech recognition
UNTIL the child says “That’s sad”, “That’s happy”, “That’s afraid”, or “That’s angry”	
The robot recognizes the picture and recalls the associated emotion	
The robot recognizes the sand-tray picture	Object (picture) recognition
The robot gives some feedback to the child	
IF child’s answer is right	
Provide very positive feedback	Say “That’s right. Well done!”
IF attempt is wrong	
Provide encouraging feedback	Say “Sorry, that’s not right. Let’s try again!”
The robot says a rational statement to the child (how to think when feeling angry, sad or scared)	
The robot recalls some associated statements from memory	
The robot says the rational statement (i.e “It is bad, but not awful when things do not happen the way you think they should happen”)	Say “<rational statement>”

The robot listens and understands what the child says
(This is not essential; the therapist could do this for the robot)
The child repeats the rational statement — Speech recognition

The robot gives some feedback to the child
(The feedback could be selected by the therapist using a remote interface)
IF child's answer is right
    Provide very positive feedback — Say "That's right. Well done!"
IF attempt is wrong
    Provide encouraging feedback — Say "Sorry, that's not right. Let's try again!"

The robot tells the child how to behave when feeling angry, sad or scared
The robot recalls some associated strategies from memory
The robot says the adaptive strategy (i.e. "When you feel angry you can breathe slowly while counting from 1 to 5") — Say "<adaptive strategy >"

Several pictures denoting different strategies appear on the sand-tray

The robot recognizes the picture
The robot recognizes the sand-tray picture — Object (picture) recognition

The robot waits until the child chooses the strategy that the robot said and shows it on the sand-tray by clicking on it

The robot recognizes the picture that the child chooses
The robot recognizes the sand-tray picture illustrating the strategies — Object (picture) recognition
Say "That's <strategy>"

The robot gives some feedback to the child
IF child's answer is right
    Provide very positive feedback — Say "That's right. Well done!"
IF attempt is wrong
    Provide encouraging feedback — Say "Sorry, that's not right. Let's try again!"

UNTIL the therapist says "Let's stop the game now" — Voice recognition (therapist)
Speech recognition

3.9 Actions to be taken when the child exhibits unexpected behaviour

The following is a simple implicit fail-safe action to be taken when a condition for continuing interaction isn't met after some pre-specified time period.

For example, in 3.8 Turn-taking Intervention, Level 2, the robot continues to monitor the head gaze direction of the child until he or she looks at the sand-tray. If, after a pre-specified period, the child doesn't look at the sand-tray then the action below will be invoked by the robot.

The same implicit strategy applies to all the other situations where the robot is waiting for the child to interact

The robot says something helpful	Say "It's okay if you don't want to do this now"
The robot says what it will do next	Say "How about playing the next game"
The robot does a little dance	Say "But let's dance first!"

The therapist selects a new task

4 Therapy Environment

Experimental Platform and System Design

Test-bed Robot

The experimental test-bed used in this study is the humanoid Nao robot developed by Aldebaran Robotics. Nao is a 25 degrees of freedom robot, equipped with an inertial sensor, two cameras (one downward-looking, one forward-looking, both with 60.9° horizontal field of view and 47.6° vertical field of view), eyes with eight full colour RGB LEDs for expressive communication, and four head-mounted microphones for spatial sound localization and automatic speech recognition.

System Design

In addition to the robot's two cameras, we use four RGB-D (colour-depth) cameras, i.e. Kinect sensors, and four high-resolution colour video cameras, all mounted remotely on a dedicated therapy table (see Section 5). Three of the RGB-D cameras are directed at the child and are used to determine the physical configuration of the child's body, e.g. the position and orientation of his or her arms. The fourth RGB-D camera is mounted above the table, facing the robot. It is used to determine the position and orientation of the robot. This is necessary because the robot has to relate the information provided by the other sensors in a real-world frame of reference to its own frame of reference, i.e. the frame of reference that is used to determine its movements.

The four colour video cameras are used to monitor the child's facial expressions, to determine the child's head gaze, and to detect, locate, and identify objects that the child may be interacting with on the table. They are also used to complement the RGB-D cameras when tracking the child's movements and monitoring her or his actions. These video cameras are placed either side of the robot, at table level and at a higher level looking down. This provides for greater flexibility when sensing the child's appearance and movements.

The final configuration, i.e. the number and layout of RGB-D sensors and cameras, will be adjusted according to sensory requirements of the interventions described in Section 3.

Environmental Setup

The studies will be conducted in a 4m x 4m testing room. The room will be split in two areas by a false wall. The left part of the room features a table and two chairs (one for the child and one for the robot). The child will interact directly with the robot that will be seated on the chair or on the therapy table. In the right part of the room the operator will control the robot's movements by using a Wizard of Oz paradigm (in our first experiments). The video cameras and sensors will be placed in the experimental room, behind the robot so as to capture the facial expressions of the children, the gaze and the movements as they interacted with the robot (see Figure 2).

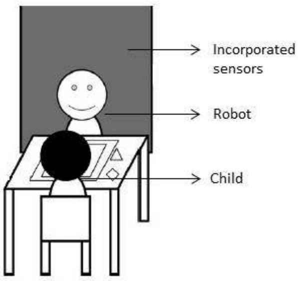

Figure 2. Description of the experimental setup.

5 Intervention Table Design

Figure 3 below shows an outline design of the therapy table to support the high-resolution video cameras, the Kinect RGB-D cameras, and the robot itself. The therapy table also provides the surface on which objects will be placed and manipulated by the robot and the child during the interventions. Note that the table will be hinged so that it can easily moved out of the way to provide space for other intervention props, such as the sand-box, and to provide sufficient space for the child to move during some of the imitation intervention exercises. A light-weight miniature gantry, comprising two uprights frames and a horizontal connecting frame, will be attached to the back of the fixed part of the table. This gantry will house the high-resolution cameras, the Kinect RGB-D cameras, and any necessary lighting. All this equipment will be camouflaged to avoid distracting the child during interventions.

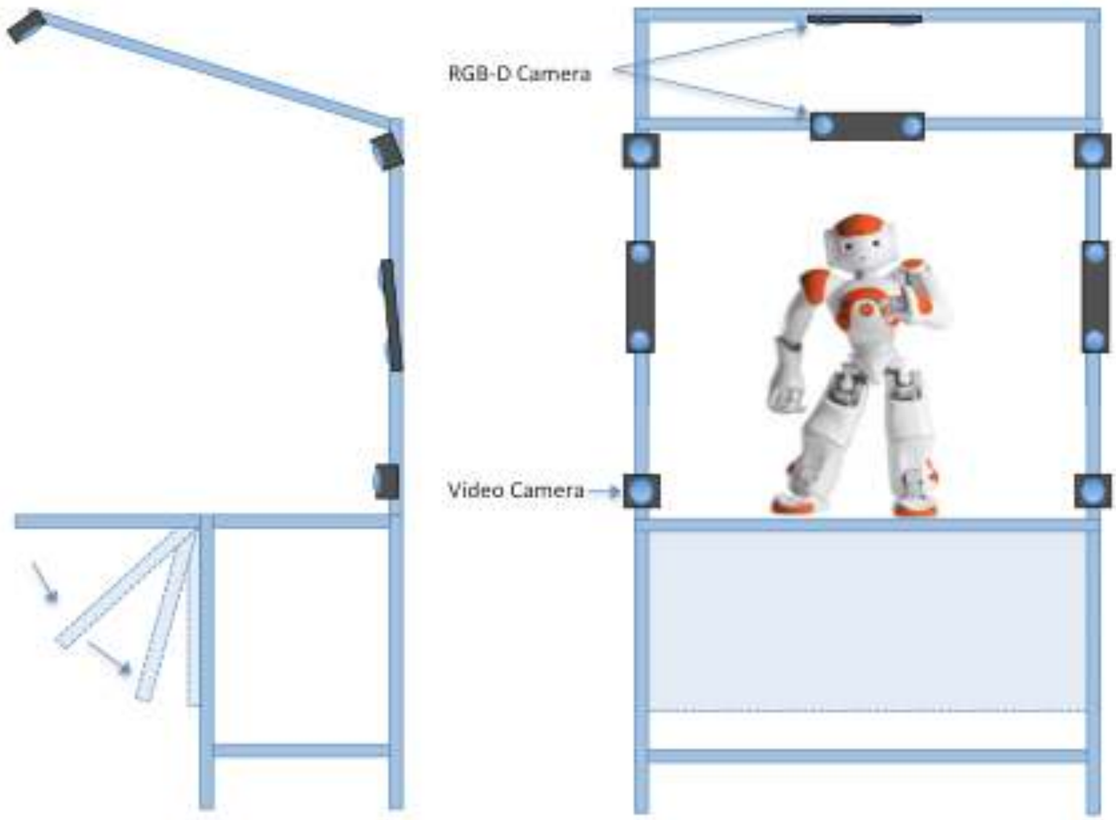

Figure 3: Schematic of the therapy table (not drawn to scale and placement of cameras for illustration only). Construction will be from light-weight modular aluminium extrusion frame components (e.g. see http://www.minitec.de/en/index.php?language=2). The downward-facing RGB-D camera (i.e. Kinect) is required to localize the NAO robot shown in the front elevation.

The dimensions of the therapy table, required for identifying the required field of view of the cameras and, hence, the focal length of the camera lenses, as well as the final CAD design, are set out in the Table 2.

Table 2: Dimensions of therapy table.

Dimension	Centimetres
Height of work surface	~ 60
Width of work surface	~ 80
Depth of foldable work surface	~ 30
Depth of fixed work surface	~ 30
Height of mounting frame	~ 140

V References

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

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. MIT press.

Baron-Cohen, S. (2009). Autism: The Empathizing–Systemizing (E-S) Theory. *Annals of the New York Academy of Sciences*, *1156*(1), 68-80.

Baron-Cohen, S., Knickmeyer, R. C., & Belmonte, M. K. (2005). Sex differences in the brain: Implications for explaining autism. *Science, 5749*(310), 819-823.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of child psychology and psychiatry*, *42*(2), 241-251.

Cardon, T. A., & Wilcox, M. J. (2011). Promoting imitation in young children with autism: A comparison of reciprocal imitation training and video modeling. *Journal of autism and developmental disorders*, *41*(5), 654-666.

Carpenter, M., Pennington, B. F., & Rogers, S. J. (2002). Interrelations among social-cognitive skills in young children with autism. *Journal of autism and developmental disorders*, *32*(2), 91-106.

Courchesne, E., Carper, R., & Akshoomoff, N. (2003). Evidence of brain overgrowth in the first year of life in autism. *Journal of the American Medical Association, 290*(3)*,* 337-344.

Courchesne, E., Karns, C. M., Davis, H. R., Ziccardi, R., Carper, R. A., Tigue, Z. D., ... & Courchesne, R. Y. (2001). Unusual brain growth patterns in early life in patients with autistic disorder: An MRI study. *Neurology, 57*(2)*,* 245-254.

Courchesne, E., Pierce, K., Schumann, C. M., Redcay, E., Buckwalter, J. A., Kennedy, D. P., & Morgan, J. (2007). Mapping early brain development in autism. *Neuron, 56*(2)*,* 399-413.

David, D., Matu, S. A., & David, O. A. (2014). Robot-Based Psychotherapy: Concepts Development, State of the Art, and New Directions. *International Journal of Cognitive Therapy*, *7*(2), 192-210.

Dawson, G., & Adams, A. (1984). Imitation and social responsiveness in autistic children. *Journal of abnormal child psychology*, *12*(2), 209-226.

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric research, 65*(6), 591-598.

Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry*, *45*(2), 212-229.

Ingersoll, B. (2011). The differential effect of three naturalistic language interventions on language use in children with autism. *Journal of Positive Behavior Interventions*, *13*(2), 109-118.

Ingersoll, B., & Schreibman, L. (2006). Teaching reciprocal imitation skills to young children with autism using a naturalistic behavioral approach: Effects on language, pretend play, and joint attention. *Journal of autism and developmental disorders*, *36*(4), 487-505.

Kozima, H., Nakagawa, C., & Yasuda, Y. (2005, August). *Interactive robots for communication-care: A case-study in autism therapy*. Paper presented at the IEEE International Workshop on Robot and Human Interactive Communication (ROMAN), Nashville, TN.

Lord, C., Risi, S., Lambrecht, L., Cook Jr., E. H., Leventhal, B. L., DiLavore, P. C., ... & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of autism and developmental disorders*, *30*(3), 205-223.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24(5),* 659-685.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). Autism diagnostic observation schedule: a standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders 19*(2)*,* 185-212.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of consulting and clinical psychology*, *55*(1), 3.

Lovaas, O. I., Freitas, L., Nelson, K., & Whalen, C. (1967). The establishment of imitation and its use for the development of complex behavior in schizophrenic children. *Behaviour Research and Therapy, 5*(3), 171-181.

Michaud, F., & Clavet, A. (2001, June). *Robotoy contest-designing mobile robotic toys for autistic children*. Paper presented at the Annual Conference of the American Society for Engineering Education, Albuquerque, NM.

Mordre, M., Groholt, B., Knudsen, A. K., Sponheim, E., Mykletun, A., & Myhre, A. M. (2012). Is long-term prognosis for pervasive developmental disorder not otherwise specified different from prognosis for autistic disorder? Findings from a 30-year follow-up study. *Journal of autism and developmental disorders*, *42*(6), 920-928.

Ozonoff, S., & Miller, J. N. (1995). Teaching theory of mind: A new approach to social skills training for individuals with autism. *Journal of Autism and developmental Disorders*, *25*(4), 415-433.

Pop., C., Pintea, S., Vanderborght, B., & David, D., (2014) Enhancing play skills, engagement and social skills in a play task in ASD children by using robot-based interventions. A pilot study. *Interaction Studies,* 15 (2), 292-320.

Ricks, D. J., & Colton, M. B. (2010, May). *Trends and considerations in robot-assisted autism therapy*. Paper presented at the IEEE International Conference on Robotics and Automation (ICRA), Anchorage, AK.

Robins, B., Amirabdollahian, F., Ji, Z., & Dautenhahn, K. (2010, September). *Tactile interaction with a humanoid robot for children with autism: A case study analysis involving user requirements and results of an initial implementation*. Paper presented at the IEEE International Workshop on Robot and Human Interactive Communication (ROMAN), Viareggio, Italy.

Rogers, S. J., Bennetto, L., McEvoy, R., & Pennington, B. F. (1996). Imitation and pantomime in high-functioning adolescents with autism spectrum disorders. *Child development*, *67*(5), 2060-2073.

Rogers, S. J., Hepburn, S. L., Stackhouse, T., & Wehner, E. (2003). Imitation performance in toddlers with autism and those with other developmental disorders. *Journal of Child Psychology and Psychiatry*, *44*(5), 763-781.

Scassellati, B., Admoni, H., & Mataric, M. (2012). Robots for use in autism research. *Annual Review of Biomedical Engineering*, *14*, 275-294.

Smith, V., Mirenda, P., & Zaidman-Zait, A. (2007). Predictors of expressive vocabulary growth in children with autism. *Journal of Speech, Language, and Hearing Research*, *50*(1), 149-160.

Stone, W. L., Ousley, O. Y., & Littleford, C. D. (1997). Motor imitation in young children with autism: What's the object?. *Journal of abnormal child psychology*, *25*(6), 475-485.

Tapus, A., Peca, A., Aly, A., Pop, C., Jisa, L., Pintea, S., ... & David, D. O. (2012). Children with autism social engagement in interaction with Nao, an imitative robot–A series of single case experiments. *Interaction studies*, *13*(3), 315-347.

Vanderborght, B., Simut, R., Saldien, J., Pop, C., Rusu, A. S., Pineta, S., ... & 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.