From Robot-Assisted Intervention to New Generation of Autism Screening: An Engineering Implementation Beyond the Technical Approach

Andrés A. Ramírez-Duque

FEDERAL UNIVERSITY OF ESPIRITO SANTO - UFES, BRAZIL
GRADUATE PROGRAM ON ELECTRICAL ENGINEERING

ANDRÉS ALBERTO RAMÍREZ DUQUE

FROM ROBOT-ASSISTED INTERVENTION TO NEW GENERATION OF AUTISM SCREENING: AN ENGINEERING IMPLEMENTATION BEYOND THE TECHNICAL APPROACH

VITÓRIA
2019

ANDRÉS ALBERTO RAMÍREZ DUQUE

FROM ROBOT-ASSISTED INTERVENTION TO NEW GENERATION OF AUTISM SCREENING: AN ENGINEERING IMPLEMENTATION BEYOND THE TECHNICAL APPROACH

A Ph.D. Thesis presented to the Graduate Program on Electrical Engineering of the Federal University of Espirito Santo - UFES, Brazil, in partial fulfillment of the requirements for the degree of Doctor in Electrical Engineering

Advisor: Prof. Dr. Anselmo Frizera Neto

Co-Advisor: Prof. Dr. Teodiano Bastos

VITÓRIA

2019

Ficha catalográfica disponibilizada pelo Sistema Integrado de
Bibliotecas - SIBI/UFES e elaborada pelo autor

Ramírez Duque, Andrés Alberto, 1987-

R173f From robot-assisted intervention to new generation of autism screening: an engineering implementation beyond the technical approach / Andrés Alberto Ramírez Duque. - 2019.

120 f. : il.

Orientador: Anselmo Frizera Neto.

Coorientador: Teodiano Bastos.

Tese (Doutorado em Engenharia Elétrica) - Universidade Federal do Espírito Santo, Centro Tecnológico.

1. Robótica. 2. Assistência. 3. Transtorno do Espectro Autista. I. Frizera Neto, Anselmo. II. Bastos, Teodiano. III. Universidade Federal do Espírito Santo. Centro Tecnológico. IV. Título.

CDU: 621.3

ANDRÉS ALBERTO RAMÍREZ DUQUE

FROM ROBOT-ASSISTED INTERVENTION TO NEW GENERATION OF AUTISM SCREENING: AN ENGINEERING IMPLEMENTATION BEYOND THE TECHNICAL APPROACH

A Ph.D. Thesis presented to the Graduate Program on Electrical Engineering of the Federal University of Espirito Santo - UFES, Brazil, in partial fulfillment of the requirements for the degree of Doctor in Electrical Engineering

Approved: July 17, 2019

EXAMINING COMMITTEE

Prof. Dr. Anselmo Frizera Neto (Advisor)
Federal University of Espirito Santo

Prof. Dr. Teodiano Bastos (Co-Advisor)
Federal University of Espirito Santo

Prof. Dr. Eduardo Rocón de Lima
Spanish National Research Council (CSIC)

Prof. Dr. Adriano de Oliveira Andrade
Federal University of Uberlândia

Prof. Dr. André Ferreira
Federal University of Espirito Santo

Prof. Dr. Eliete Maria de Oliveira Caldeira
Federal University of Espirito Santo

I would like to dedicate my work to my wife Ximena and my children Pablo and Luciana, without whom this effort would not have been possible. Also, I dedicate this thesis to my parents Gloria Emilia and Luis Alberto and to all people who guided and supported me in this process.

ACKNOWLEDGEMENTS

First of all, I will start by thanking Prof. Dr. Anselmo Frizera-Neto for agreeing to guide me in this process. Although the research area was new to him, without hesitation, he decided to support me. Anselmo was an unbeatable advisor, always checking my manuscripts, and clarifying my doubts. Anselmo understands my virtues and my weaknesses and as a good leader helped me to improve the right thing and to deal with wrong thing. Besides, Anselmo gave me his knowledge and experience, but also he granted us, my family and me, with his friendship and unconditional support. We always will grateful, thank you!

Secondly, I would like to thanks Prof. Dr. Teodiano Bastos for his kindness and disposition, for his expert and accurate advices. I am grateful for the time you took to read all my manuscripts, letter, and emails, as well as for all the coffee time you had to listen to me.

As important are my wife and children that became my biggest motivation throughout this process. Thanks for every laugh, tear and mischief. I love you!.

I want to acknowledge the children that participated in my studies together with their families. Thank you for your trust, interest, opinions, and perspectives to make every intervention-day possible. I am grateful to Howard Gardner Clinic and the Colombian School of Engineering Julio Garavito, who offered me the support to work in Colombia.

I would like to thank specially to Mario Jimenez and Rosa Muñoz for all words of encouragement in difficult days and so many anecdotes built together. Gracias mijo!

I thanks to all my colleagues and classmates of the NTA group from UFES/Brazil, who shared with me a lot of quality life experiences.

This work was supported in part by the Royal Academy of Engineering, CASTOR Project: CompliAnt SofT Robotics (grant IAPP1\100126), and the first author scholarship was supported in part by the *Coordenãço de Aperfeiçoamento de Pessoal de Nivel Superior* - Brasil (CAPES) - Finance Code 001.

ABSTRACT

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition that affects people from birth, whose symptoms are found in the early developmental period. The ASD diagnosis is usually performed through several sessions of behavioral observation, exhaustive screening, and manual coding behavior. The early detection of ASD signs in naturalistic behavioral observation may be improved through Social Assistive Robotics (SAR) and technological-based tools for an automated behavior assessment. Robot-assisted tools using Child-Robot Interaction (CRI) theories have been of interest in intervention for children with Autism Spectrum Disorder (CwASD), elucidating faster and more significant gains from diagnosis and therapeutic intervention when compared with classical methods. Additionally, using computer vision to analyze the child's behavior and automated video coding to summarize the responses would help clinicians to reduce the delay of ASD diagnosis.

Despite the increment of researches related to SAR, achieving a plausible Robot-Assisted Diagnosis (RAD) for CwASD remains a considerable challenge to the clinical and robotics community. The work of specialists regarding ASD diagnosis is hard and labor-intensive, as the condition's manifestations are inherently heterogeneous and make the process more difficult. In addition, the aforementioned complexity may be the main reason for the slow progress in the development of SAR with diagnostic purpose. Also, there still is a lack of guidelines on how to select the appropriate robotic features, such as appearance, morphology, autonomy level, and how to design and implement the robot's role in the CRI.

Thus, this Ph.D. Thesis provides a comprehensive Robot-Assisted intervention for CwASD to assess autism risk factors for an autism diagnostic purpose. More specifically, two studies were conducted to analyze and validate the system performance. Through statistical data analysis, different behavior pattern of the CwASD group were identified, which suggest that these patterns can be used to detect autism risk factors through robot-based interventions. To increase the scope of this research, a theoretical conceptualization of the pervasive version of the multimodal environment was described as well as a participatory design methodology was designed and implemented on the Colombian autism community, providing, a set of guidelines regarding the design of a social robot-device suitable to be applied for robot-assisted intervention for CwASD.

Keywords: Autism Spectrum Disorder. Social Assistive Robotics. Child-Robot Interaction. Computer Vision Systems. Participatory Design.

List of Figures

List of Tables

LIST OF ABBREVIATIONS AND ACRONYMS

ADHD	Attention Deficit Hyperactivity Disorder
ASD	Autism Spectrum Disorder
AU	Action Units
CDC	Centers for Disease Control and Prevention
CE-CLM	Experts Constrained Local Model
CG	Control Group
CHAT	Checklist for Autism in Toddlers
CLM	Constrained Local Model
CLNF	Conditional Local Neural Fields
CNN	Convolutional Neural Network
CRI	Child-Robot Interaction
CwASD	Children with ASD
CwDC	Children with Different Condition
DOF	Degrees of Freedom
DSM-V	Diagnostic and Statistical Manual of Mental Disorders, 5th edition
DTT	Discrete Trial Teaching
FACS	Facial Action Coding System
GARS	Gilliam Autism Rating Scale
HG	Howard Gardner
HMM	Hidden Markov Model
HRI	Human-Robot Interaction
ICC	Intraclass Correlation Coefficient
IJA	Initiation of Joint Attention

IO-HMM	Input-Output Hidden Markov Model
IRB	Initiation of Request Behavior
JA	Joint Attention
LNF	Local Neural Field
M-CHAT	Modified Checklist for Autism in Toddlers
MERI	Multimodal Environment for Robot-mediated Intervention
MERI-AI	MERI Annotator Interface
MERI-PI	MERI Protocol Interface
OPSORO	Open Source Platform for Social Robotics
PD	Participatory Design
PDM	Point Distribution Model
PHS	Pervasive Healthcare System
RAD	Robot-Assisted Diagnosis
RAT	Robot-Assisted Therapy
RET	Robot-Enhanced Therapy
RJA	Responding to a bid for Joint Attention
RMI	Robot-Mediated Intervention
ROS	Robot Operating System
SAR	Social Assistive Robotics
TD	Typically developing
TMI	Therapist-Mediated Intervention
UVV	University of Vila Velha
VFOA	Visual Focus of Attention

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Autism Spectrum Disorder (ASD)

According to Centers for Disease Control and Prevention (CDC) of USA, Autism Spectrum Disorder (ASD) is a neuro-developmental disorder characterized by a widespread range of conditions related to some degree of impaired social interaction, communication, and language, as well as severely restricted interests and highly repetitive behavior (ASSOCIATION, 2013). ASD is considered as a pervasive disorder, as it affects people from birth and tends to persist into adolescence and adulthood. In most cases, the conditions are apparent from 38 to 48 months of life (STEINER *et al.*, 2012). The diagnostic criteria for ASD included in the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-V)(ASSOCIATION, 2013), refer to ASD as a single diagnosis category that includes autistic disorder (autism), Asperger's disorder, childhood disintegrative disorder, and pervasive developmental disorder not otherwise specified (ASSOCIATION, 2013).

Children with ASD (CwASD) often have different forms of language acquisition and information processing, such as the ability to read social cues, understand social interaction, and respond appropriately (CARVALHO *et al.*, 2014). In addition, they can present a large variability in attention, challenging to understand the concept of time, and higher preference to pre-established routines. Also, individuals with ASD exhibit behaviors that include compulsions, echolalia, and motor mannerisms, such as hand flapping and body rocking (MATSON; RIVET, 2008), and other co-occurring conditions, such as epilepsy, depression, anxiety, Attention Deficit Hyperactivity Disorder (ADHD), can be manifested (ASSOCIATION, 2013). Many of these individuals can live independently, but those with severe conditions degree require life-long care and support.

The term *spectrum* of ASD is used to denote the large heterogeneity in the condition's manifestation and the range of how the aforementioned symptoms impact kids. That impact is described by a support level among 1-3 where 3 indicates that a high level of support will be needed (ASSOCIATION, 2013). Thus, autism understanding models, as well as the research regarding therapy and diagnosis, are continuously evolving. Currently, there is an active campaign to remember that autism cannot be understood as a disorder fixed in time or space, which has generated different perspectives on the condition of autism. For example, in (RICHARDSON *et al.*, 2018) proposed a ethical-based model of autism, affirming that individuals with autism have a different sociability, rather than an absent one, i.e., children with autism can exhibit strong emotional attachments to their primary caregivers, and express interest in other activities, besides engaging with technological items. They also enjoy relationships with animals, physical activity, and artistic play.

According to the CDC of USA, ASD occurs in 1.42% of the population and is almost five times more common among boys than girls (NUNES; WALTER, 2018). In developing nations, as Brazil, this condition seems to affect 0.3% of the people, as indicated in a study conducted by (PAULA *et al.*, 2011). In Colombia, the Colombian League of Autism estimates that 0.01% of people are diagnosed with ASD [1].

1.1.1 Development of Autism and Joint Attention

As CwASD have a different understanding of social behaviors, three domains have been identified as crucial to improve their relationship with others (BEAN; EIGSTI, 2012): social interaction, communication, and learning. Thus, turn-taking, Joint Attention (JA), and imitation tasks have been chosen as pivotal activities as through these tasks the previous domains can be trained (HUIJNEN *et al.*, 2016a). Turn-taking involves reciprocal interaction with peers, and is essential for mutual learning. Imitation is a vital human skill for social cognition and helps supporting verbal and nonverbal communication, as well as, cognitive development (DAWSON *et al.*, 2004). JA is the ability to attend to objects in the same space, and is determined through pointing or gaze gestures (CHARMAN, 2003).

In particular, JA is considered a more complex construction of the social process that

[1] Liga Colombiana de Autismo (https://www.ligautismo.org).

involves sharing an interest with another person on an object or event (MOORE; DUNHAM, 1995). JA can be divided into three classes: Initiation of Joint Attention (IJA), Responding to a bid for Joint Attention (RJA) and Initiation of Request Behavior (IRB) (ZAQUEU *et al.*, 2015). In the literature, researchers affirm that engaging in JA precede pro-social cooperative behaviors, such as social-communicative functioning, language development, empathy, and theory of mind (CHARMAN, 2003). The early exhibition of JA is strongly associated with later cognitive ability, and atypical development of IJA is strongly indicative of ASD (MOORE; DUNHAM, 1995). JA can be perceived from the first ten months of life and follows a trajectory through toddlerhood, and continues as children acquire more sophisticated social and linguistic abilities (DAWSON *et al.*, 2004).

1.1.2 Assessment and Diagnosis of ASD

ASD diagnosis is a complex process composed of several steps, involving many people interpreting intricate information from multiple sources regarding a child's behavior (WIGGINS *et al.*, 2006). The diagnosis road map begins with the first concern, which appears when parents or caregivers recognize unusual child's behaviors. Children exhibit the first signs of ASD between six and twelve months, becoming more noticeable and stable between 18 and 24 (WIGGINS *et al.*, 2015). ASD signs that may be noticed early are impairment in eye contact and shared attention, reduced social interest and in pleasure-sharing games, as well as other global signs of delay (JOHNSON; MYERS, 2007).

Before the first concern, trained pediatricians can recognize the signs and symptoms of ASD and may begin a systematic assessment; this stage is called identification of risk factors. At this point, screening tools are used through observational evaluation to recognize and confirm these signs and symptoms of (VALICENTI-MCDERMOTT *et al.*, 2012). Finally, standardized tools for the comprehensive diagnosis are applied for addressing the severity level of the ASD. Some examples of such devices are the Checklist for Autism in Toddlers (CHAT) and the Modified Checklist for Autism in Toddlers (M-CHAT) (ROBINS; DUMONT-MATHIEU, 2006). According to (BRETT *et al.*, 2016), the average age of autism diagnosis is 55 months.

1.2 Currently technology trends in ASD interventions: therapy and diagnosis

Social Assistive Robotics (SAR) is an established research area in robotics in which robots are used to support individuals through social interactions rather than physical interactions during a range of therapeutic and healthcare interventions (BELPAEME *et al.*, 2012; SCASSELLATI *et al.*, 2012; VALLÈS-PERIS *et al.*, 2018). Promising results exist in therapeutic interventions for children, elderly, stroke patients, and special-needs populations (Robotics Technology Consortium, 2013). The introduction of social robotics in real-world healthcare practices is proof of a change in people's attitudes towards the application of robotics in general and Human-Robot Interaction (HRI) in specific.

One of the most valuable contributions of SAR has been the support for ASD interventions. In the context of ASD therapy, SAR has been used to practice social skills (CABIBIHAN *et al.*, 2013). Skills such as making eye contact and recognizing emotions (YUN *et al.*, 2017), joint attention (SCASSELLATI *et al.*, 2018; KUMAZAKI *et al.*, 2018b), increasing self-initiated interactions (DAVID *et al.*, 2018), and sharing in simple activities, to encourage basic verbal and non-verbal communication (KIM *et al.*, 2013).

SAR applications for child population have promoted the growth of new theories related to the HRI paradigm known as Child-Robot Interaction (CRI). Thus, CRI researches aim to provide the necessary conditions for the interaction between a child and a robotic device taking into account some key features, such as child's neurophysical and physical condition, and the child's mental health (BELPAEME *et al.*, 2013). That is how CRI theories have been used to build robot-based interventions for CwASD, elucidating faster and more significant gains from the therapeutic intervention when compared with traditional therapies (SCASSELLATI *et al.*, 2012; PENNISI *et al.*, 2016; SCASSELLATI *et al.*, 2018; KUMAZAKI *et al.*, 2018a).

In SAR interventions for CwASD, the robot is used as an alternative communication channel to elicit a particular social response. The robotic system can also provide quantitative measurements of behaviors that may be used to characterize the child's condition, as through the use of sensors, the robot can record responses and estimate quantitative measurements (SCASSELLATI, 2007). Thus, new SAR interventions are being equipped with computer vision systems to automatically monitoring different child's behaviors.

Automated behavior evaluation systems and automatic video coding to summarize the interventions can help clinicians to reduce the delay of ASD diagnosis, decrease their workload, and improve the therapy practices, allowing to provide the CwASD with access to early and enhanced therapeutic interventions (BELPAEME *et al.*, 2012).

1.2.1 Motivation behind of using robotics systems as assistive tool in ASD intervention

Recent researches have shown the acceptance and efficiency of technologies used as auxiliary tools for intervention with ASD population (HUIJNEN *et al.*, 2016a; ARESTI-BARTOLOME; GARCIA-ZAPIRAIN, 2014; BOUCENNA *et al.*, 2014; GRYNSZPAN *et al.*, 2014). Such techniques may also be useful for community surrounding ASD individuals (therapists, caregivers, family members). For example, the use of artificial vision systems to measure and analyze the child's behavior can lead to alternative screening and monitoring tools that help clinicians to get feedback from the effectiveness of the intervention (REHG *et al.*, 2014). Additionally, social robots have great potential for aid in the diagnosis and therapy of CwASD (HUIJNEN *et al.*, 2016a; CABIBIHAN *et al.*, 2013). In fact, higher degree of control, prediction, and simplicity may be achieved in interactions with robots, impacting directly on frustration and reducing the anxiety of these individuals (SARTORATO *et al.*, 2017).

Several research have hypothesized and suggested that if social information is presented in a simplified, ASD individuals can easily understand and identify the expected behavior and engage more in social interaction (SCASSELLATI *et al.*, 2012; BELPAEME *et al.*, 2012; WAINER *et al.*, 2014). In other words, people transmit nonverbal messages with their facial expression, body posture, and voice tone (whether or not consistent with verbal information), but making sense of nonverbal communication is challenging for CwASD. In contrast, robots do not have these nonverbal elements and ambiguity in their communication, thus, interacting and learning with robots may be more comfortable and more pleasant for CwASD (CABIBIHAN *et al.*, 2013; PENNISI *et al.*, 2016).

Robot-based intervention studies have shown significant progress since many positive outcomes were found in its use as an assistive tool in therapy and diagnosis (COSTESCU *et al.*, 2014; SCASSELLATI *et al.*, 2018). Pennisi *et al.* reported that CwASD often perform tasks better with a robot partner than with a human partner. The authors also state that CwASD present similar behavior towards robots that Typically developing (TD) children present towards humans (PENNISI *et al.*, 2016). Furthermore, CwASD show a reduction in repetitive patterns of behavior and an increase in the spontaneous pro-social behaviors during therapy sessions with robots (ESTEBAN *et al.*, 2017). These findings suggest that social robots have the potential to be used as assistive tools for evidence-based intervention.

Respect to the use of computer vision techniques, previous studies already analyzed child's behaviors, such as visual attention, eye gaze, eye contact, smile events, and visual exploration using cameras and eye trackers (CHONG *et al.*, 2017; NESS *et al.*, 2017), and RGBd cameras (REHG *et al.*, 2013; CAZZATO *et al.*, 2016). These studies have shown the potential of vision systems in improving the behavioral coding in ASD therapies. Currently more elaborated system implement automated face analysis and artificial cognition through robot-mediator, which analyze child's engagement (LEMAIGNAN *et al.*, 2016; COCO *et al.*, 2017), emotions recognition capability (LEO *et al.*, 2015; PALESTRA *et al.*, 2016; POUR *et al.*, 2018), and child's intentions (ESTEBAN *et al.*, 2017; FENG *et al.*, 2018). In these works, two different strategies were implemented, where the most common is based on mono-camera approach using an external RGB or RGBd sensor (LEMAIGNAN *et al.*, 2016; COCO *et al.*, 2017; POUR *et al.*, 2018) or using on-board RGB cameras mounted on the robotic-platform (LEO *et al.*, 2015; FENG *et al.*, 2018). Other strategies are based on a highly structured environment composed of an external camera plus an on-board camera (PALESTRA *et al.*, 2016) or a network of vision sensors attached to a small table (ESTEBAN *et al.*, 2017). These strategies based on multi-camera methods improve the performance of the sensing system, but remain constrained in relation to some desired features, such as flexibility, scalability, and modularity.

1.2.2 Participatory design and novel approaches to design robot mediated interventions

Currently, there is still a lack of consensus on how the robot-mediated intervention should be addressed, and which robot's role and robot morphology might be most effective. Most robots used with ASD populations are off-the-shelf robots, from toy robots to social robots, which were not designed explicitly for the ASD population or therapeutic or diagnostic ASD intervention (VALLÈS-PERIS *et al.*, 2018).

Several design techniques have been explored over the years, which all integrate contributions from different populations affected by the design decisions (e.g., stakeholders community) (FLETCHER-WATSON *et al.*, 2018). The Participatory Design (PD) process is a well-known strategy in industrial design and the arts to develop products and services for a target population. The philosophy behind PD is to empower the people that are involved in a specific activity or situation by providing them with space and a voice so that all can contribute in the decision making (GUHA *et al.*, 2014). The process intends to, in the end, achieve products or services that represent the real needs, desires, and expectations of the users, designers, and stakeholders. The application of PD techniques is particularly promising when transferring knowledge and systems from research to the real world, primarily if the success of the product or service hinges on the interaction with the human.

The use of PD methods in the design of technology-based processes for health care is a recognition of both users and stakeholders as “experts” in their fields, highlighting the different experiences and attitudes that they may have (FLETCHER-WATSON *et al.*, 2018). In this sense, all the actors in the project are recognized as valuable contributors, which plays a crucial role in ethical, political, and social considerations of the development. The target populations and their environment (families, society, groups of allies and friends) are no longer seen as a source to obtain information and requirements to produce results, but rather a partner with experience and a different way to see the world that can be a part of the solution (MERTER; HASIRCI, 2016).

PD has been used in the design of SAR for ASD (HUIJNEN *et al.*, 2016a) and development of HRI in the healthcare systems (VALLÈS-PERIS *et al.*, 2018). However, implementing a participatory or co-design process with ASD populations can be very challenging. Researchers and designers need to find ways and techniques to overcome several limitations of traditional co-creation methods as they have to establish additional *modus operandi* to choose and adapt co-design techniques based on their participants’ abilities (i.e., their strengths and skills) rather than their disabilities.

1.3 Objectives and outline of this Ph.D. Thesis

By adopting a new engineering perspective beyond the technical approach, this doctoral thesis aims to explore whether there is a potential of robot-based intervention for autism screening, by going beyond weaknesses in contemporary SAR and CRI engineering applications, and implementing technological tools in an ethical way that will help to examine the utility of robots in ASD diagnosis clearly. Thus, the global objective of this research is to explore the usefulness of a multimodal environment for robot-mediated intervention as an alternative tool to assess autism risk factors associated with JA behaviors. The following specific objectives are proposed to reach the research purpose:

1. To develop a multimodal environment for robot-mediated intervention, able to provide social prompts for CwASD.
2. To design a child-robot interaction protocol to assess joint attention behaviors exploring clinical aspects and evidence-based practice guidelines.

3. To develop studies with children with special needs to validate the usefulness of the system as an assistive tool for the autism diagnosis.
4. To collect and analyze the autism community perspective regarding child-robot interaction scenarios to enhance robot-mediated practices.

This thesis begins with an overview of SAR for autism diagnosis and robot-mediated intervention used in JA assessment in Chapter 2, where the underlying SAR application and results found in the literature, specifically on JA interventions for both, ASD therapeutic and diagnostic purpose are explained. Chapter 3 provides a detailed explanation of the multimodal environment for robot-based intervention, where several contributions were developed.

Two studies developed to analyze and validate the system performance are presented in Chapter 4, in addition to statistical analysis and main findings regarding a comparative assessment of JA performance conducted in a Colombian clinic. In Chapter 5, a theoretical conceptualization of the pervasive version of the multimodal environment is presented, as well as, an alternative and feasible three-step process to identify the pathways from the first concern to pervasive ASD diagnosis.

Chapter 6 exhibits the design, implementation, and results of a participatory design method applied in Colombia with and for an Autism community. Finally, Chapter 7 provides the concluding remarks of this thesis and outlines some general discussion together future research directions.

1.4 Publications

The research developed in this Ph.D. thesis allowed the publication of the following works:

1. (Journal Paper) Andrés A. Ramírez-Duque, Anselmo Frizera, Teodiano Bastos. Robot-Assisted Autism Spectrum Disorder Diagnostic Based on Artificial Reasoning. Journal of Intelligent and Robotic Systems. March, 2019. DOI: 10.1007/s10846-018-00975-y.
2. (Conference Proceeding) Andrés A. Ramírez-Duque, Anselmo Frizera and Teodiano Bastos. Robot-Assisted Diagnosis for Children with Autism Spectrum Disorder Based on Automated Analysis of Nonverbal Cues. 7th IEEE International Conference on Biomedical Robotics and Biomechatronics - Biorob2018, Enschede, Holland. ISBN: 978-1-5386-8183-1, 2018

3. (Conference Proceeding) Andrés A. Ramírez-Duque, Anselmo Frizera and Teodiano Bastos. Sistema para identificar déficit de atención compartida en niños con trastorno del espectro autista a partir de la estimación del foco de atención visual por red de sensores RGB y RGBd. II Congreso Internacional de Tecnología y Turismo para Todas las Personas, 2017, Málaga. p. 125-134.
4. (Conference Proceeding) Andrés A. Ramírez-Duque, Anselmo Frizera and Teodiano Bastos. Estimación automática del foco de atención visual para identificar factores de riesgo en niños con trastorno del espectro autista. IX Congreso Iberoamericano de Tecnología de Apoyo a la Discapacidad, 2017, Bogotá. p. 461-468.

Other works were also published as a consequence of the interaction with other researchers during the development of this Ph.D. thesis. The most important ones are listed below.

1. (Journal Paper) Nicolas Valencia-Jimenez, Arnaldo Leal-Junior, Leticia Avellar, Laura Vargas-Valencia, Pablo Caicedo-Rodrigues, Andrés A. Ramirez-Duque, Mariana Lyra, Carlos Marques, Teodiano Bastos, Anselmo Frizera. A Comparative Study of Markerless Systems Based on Color-Depth Cameras, Polymer Optical Fiber Curvature Sensors, and Inertial Measurement Units: Towards Increasing the Accuracy in Joint Angle Estimation. ELECTRONICS, v. 8, p. 173, 2019.
2. (Conference Proceeding) Thiago L. Carvalho, Andrés A. Ramírez-Duque, Anselmo Frizera, Teodiano Bastos. Estudo de precisão de uma platadorma multi câmeras RGBd para sistemas de reabilitação. V Congresso Brasileiro de Eletromiografia e Cinesiologia e X Simpósio de Engenharia Biomédica, 2018, Uberlândia. p. 516-519.
3. (Conference Proceeding) Thiago L. Carvalho, Andrés A. Ramírez-Duque, Anselmo Frizera, Teodiano Bastos. Sistema Multi-Câmeras RGBd para Cálculo de Parâmetros Espaço-Temporais da Marcha Humana. IX Congreso Iberoamericano de Tecnología de Apoyo a la Discapacidad, 2017, Bogotá. p. 40-47.
4. (Conference Proceeding) Camila R. Carvalho, Carolina C. Carvalho, Andrés A. Ramírez-Duque, Anselmo Frizera-Neto. Adaptação de um sistema robótico para comunicação e assistência no diagnóstico de crianças com Transtorno do Espectro Autista. IX Congreso Iberoamericano de Tecnología de Apoyo a la Discapacidad, 2017, Bogotá. p. 505-512.

In addition, the following manuscripts were submitted to journals:

1. (Journal Paper) Andrés A. Ramírez-Duque, Teodiano Bastos, Marcela Munera, Carlos Cifuentes, Anselmo Frizera-Neto, (2019). Robot-Assisted Intervention for Children with Special Needs: A Comparative Assessment for Autism Screening. *Manuscript submitted for publication*.
2. (Journal Paper) Andrés A. Ramírez-Duque, Luis Aycardi, Adriana Villa, Marcela Munera, Teodiano Bastos,Tony Belpaeme, Anselmo Frizera-Neto, Carlos Cifuentes, (2019). Collaborative and Inclusive Process with the Autism Community: A Case Study of Social Robot Design. *Manuscript submitted for publication.*
3. (Journal Paper) Andrés A. Ramírez-Duque, Anselmo Frizera, Teodiano Bastos, (2019). Multimodal Environment for Robot-Mediated Intervention (MERI): A New Paradigm for Pervasive ASD Diagnosis. *Manuscript submitted for publication*

CHAPTER 2

ROBOTICS SYSTEMS IN ASD DIAGNOSIS INTERVENTION

Currently, Social Assistive Robotics and Child-Robot Interaction researches have shown prominent results that have aroused a lot of interest within the autism community (SCASSELLATI *et al.*, 2018). However, achieving a plausible Robot-Assisted Diagnosis (RAD) for Children with Autism Spectrum Disorders remains a considerable challenge to the clinical and robotic community. Surprisingly, the view of assisting the autism diagnosis using robots is not new. In 2007, (SCASSELLATI, 2007) proposed that robot-assisted diagnosis would be one of the applications with the most significant potential for the autism community. However, there are few reported works regarding robot-based diagnostic tools. In contrast, many works regarding Robot-Assisted Therapy (RAT) and an improved version called Robot-Enhanced Therapy (RET) have been reported in (ESTEBAN *et al.*, 2017). Currently, the development of RAT applications continues increasing. Example of this trend is reflected in the reviews on (PENNISI *et al.*, 2016), which the authors reported 25 RAT from 2009 to 2016. Also, in (BEGUM *et al.*, 2016) they reported a survey with 14 RAT studies between 2009 and 2014. However, no studies regarding assisted diagnosis tools were found in these reviews due to there are few studies.

The slow progress in the development of SAR-based tools to help specialists in the diagnosis of autism may have been the cause of the lack of these studies (SCASSELLATI, 2007). In fact, the traditional diagnosis of autism is already stressful enough, given the high variability of signs exhibited for CwASD (STEINER *et al.*, 2012). The traditional process of ASD diagnosis requires from the medical specialist to address behavioral assessments of the child's development state in four domains, such as behavior excesses, communication, self-care, and social skills (WIGGINS *et al.*, 2015). These assessments require responses to a large number of paper-based questionnaires, which makes many of them lengthy and labor-intensive, such as CHAT (BARON-COHEN *et al.*, 1992), and Gilliam Autism Rating Scale (GARS) (LECAVALIER, 2005). In addition, the literature shows that there is a time gap among parents' first concern about the child's development impairments, their first medical evaluation, and the child's age of confirmed diagnosis (HUERTA; LORD, 2012). Furthermore, it has been identified that child's care and education centers have more opportunities to recognize risk factors of ASD in children than pediatric surveillance system (WIGGINS *et al.*, 2006).

From a technical perspective, the lag in manifesting real benefits of robots to make autism screening is also due to the difficulty of finding mechanisms to analyze children in naturalistic environments, elicit behaviors through few robot interventions and also generate clinically valid metrics that allow identifying behavioral patterns to confirm risk factors of ASD (SCASSELLATI, 2007).

Although the concerns and objectives associated with RAT and RAD interventions with CwASD, in general, are different, many of the RAT's advances and positive findings can be directly or indirectly applied in the development of applications for autism screening using a robotic-assistive approach. For example, the already clear need to use well-standardized clinical protocols, the use of good and rigorous experimental designs to generate clinical evidence (BEGUM *et al.*, 2016), and the findings regarding the acceptability of robot's physical features and automated behaviors (SIMUT *et al.*, 2016; CABIBIHAN *et al.*, 2013) are all applicable to improve the screening processes and identification of risk factors using the paradigm of children's interaction with a robot-mediator. Thus, RAT researches have shown that CRI benefits are promising in therapies regarding behavioral domains such as joint attention, imitation task, and free play (PENNISI *et al.*, 2016; HUIJNEN *et al.*, 2016a; HUIJNEN *et al.*, 2016b). The child's response towards a robotic platform in these domains also can be used as diagnostic inputs (SCASSELLATI *et al.*, 2012).

2.1 Robot-based intervention for autism screening

The justification in (TAPUS *et al.*, 2007) for using robots as autism screening tools continues to be valid. A robotic system is understood as an interactive platform composed of many subsystems with the ability to obtain information about the children-robot interaction passively as well as actively and to provide systemically social cues designed to elicit particular social responses (SCASSELLATI, 2007). The use of a robotic platform allows creating systematic social prompts to stimulate children and generate response composed of specific social behaviors. Thus, robot-based interactions help to find unique opportunities in which quantitative quality metrics about pattern behaviors can be analyzed. These metrics provide a chance to compare children in a standardized manner as well as the possibility of assessing the progress of a single individual across time and contexts. One possible outcome of this approach is a robot-based screening technique capable of detecting risk factors for autism in infants and toddlers.

Robot-assisted diagnosis leads to an understanding of the social responses of children in a context-dependent environment that facilitates both the acquisition and analysis of evidence. For example, the robotic system can generate stimuli in a repeatable and standardized way to elicit children behaviors that may not emerge in traditional diagnostic process. In addition, the system can record evidences autonomously and compare these throughout sessions, different environments and scenarios (KIM *et al.*, 2012). From the clinician point of view, observing the child's behavior in a multiple environment is considered the most suitable method to diagnose autism. In fact, the comprehensive assessment done in the context of daily life can lead to an accurate ASD diagnosis (HUERTA; LORD, 2012).

In the line of (SCASSELLATI, 2007) and (TAPUS *et al.*, 2007) researches, in (DEHKORDI *et al.*, 2015) the authors developed a parrot-like robot called Robot-Parrot for screening autistic children. They used 12 diagnostic criteria that are listed in the DSM-V as a standard protocol to assess children's behavior. In (MOGHADAS; MORADI, 2018) the authors extended the work of (DEHKORDI *et al.*, 2015) automating the process of interaction with a robot parrot to distinguish between TD children and CwASD.

2.2 Joint attention assessment using robot-assisted intervention

Researchers of Vanderbilt University published a series of studies showing an experimental protocol to assess JA. The protocol consisted of directing the attention of the child towards objects located in the room through adaptive prompts (BEKELE *et al.*, 2013). In (BEKELE *et al.*, 2014) the authors inferred the participant's eye gaze by the head pose, which was calculated in real-time by an IR camera array . In their last works (ZHENG *et al.*, 2013; WARREN *et al.*, 2015), they used a commercial eye tracker to estimated the children's eye gaze around the robot and manual behavioral coding for global evaluation. However, eye tracker devices require pre-calibration and may limit the movement of the individual. The results of these works showed that the robot attracted children's attention and that CwASD reached all JA task. Nevertheless, developing JA tasks is more difficult with a robot than with humans (WARREN *et al.*, 2015). In (ANZALONE *et al.*, 2014) the authors developed a CRI scenario using the NAO robot to perform JA tasks, in which they used an RGBD camera to estimate only body and head movements. The results showed that JA performance of children with ASD was similar to the performance of TD children when interacting with the human mediator, however, with a robot mediator, the children with ASD presented a lower performance than the TD children, i.e, the children with ASD needed more social cues to finalize the task. In (CHEVALIER *et al.*, 2016) they analyzed some features, such as proprioceptive and visual integration in CwASD, using an RGBD sensor to record the interventions sessions and manual behavior coding to analyzed the participants' performance. In none of the previous works, a closed-loop subsystem was implemented to provide some level of artificial cognition to enable automated robot behavior. All of aforementioned works were reviewed and analyzed in (PENNISI *et al.*, 2016), they concluded that, in a general perspective more and best structured studies are required to gather more evidences and valid clinical findings.

Afterwards, between 2017 and 2018, seven more studies have been reported (see Table 1). In order to identify practices and techniques as well as expected outcomes that can also be extended to RAD, an analysis of the most representative methodological and technical characteristics of those works is presented below.

Four works used the robot NAO despite its physical limitation to represent gaze shifting (PENNISI *et al.*, 2016). NAO remains the most used robotic platform to direct JA tasks. In other three studies, Jibo (SCASSELLATI *et al.*, 2018), iRobiQ together CARO (YUN *et al.*, 2017) and CommU (KUMAZAKI *et al.*, 2018a) were used to interact with children. The CommU robot was equipped with physical eyes to perform gaze shift, while Jibo was provided with animated screen-based eyes. In line with previous works, the most common robot control strategy was the *Wizard of Oz*, that means, teleoperation control. However, one work showed semi-autonomous control and two used an autonomous control approach. Moreover, despite that the seven researches used vision sensor arrays, only in (CAI *et al.*, 2018) the authors implemented algorithms to automatically estimate performance metrics using the data provided for the vision system.

Regarding the implemented protocols, the seven papers chose protocols used in traditional therapies concerning JA training composed of different prompt levels administered progressively and sequentially to conduct the child' attention towards different objects or events. In addition, the most used prompts were looking, pointing and talking and, there is a consensus that a combination of these prompts increases child's performance in the intervention. The metrics used to estimate the child's performance were JA score and eye contact. Furthermore, in (ANZALONE *et al.*, 2018) the authors also proposed alternative metrics, such as child's energy and displacement throughout the session.

The experimental design implemented was different in each research. In general, single-case with intensive therapies administered for multiple session was the most common design (ZHENG *et al.*, 2018; DAVID *et al.*, 2018; SCASSELLATI *et al.*, 2018; ANZALONE *et al.*, 2018), and only two cases implemented a group-based model (KUMAZAKI *et al.*, 2018a; YUN *et al.*, 2017). A description of the different technical and clinical aspects of these works is summarized in Table 1.

Finally, the conclusions of the above works showed that there is still no consensus regarding the effects of the mediator and regarding the differences in behavior between the CwASD and TD children. For example, in (YUN *et al.*, 2017) the authors stated that there is no significant difference between the performance obtained with the robot mediator compared with the score obtained in the scenario of the human mediator. On the contrary, in (KUMAZAKI *et al.*, 2018a) they found that CwASD were better during the robotic intervention than during the human agent intervention. On the other hand, in (ANZALONE *et al.*, 2018) they affirmed that TD children respond more than CwASD to the JA induction performed by the robot while in (KUMAZAKI *et al.*, 2018a) the authors say that the effect of the robot mediator was lower in TD than in ASD.

	(ANZALONE *et al.*, 2018)	(CAI *et al.*, 2018)	(DAVID *et al.*, 2018)	(KUMAZAKI *et al.*, 2018a)	(SCASSELLATI *et al.*, 2018)	(YUN *et al.*, 2017)	(ZHENG *et al.*, 2018)
Goals	Explored the dynamics of JA in children with Autism ASD during an interaction task with a small humanoid robot	Improved the systems of both standard and robot assisted therapy for CwASD via a sensing framework with multisensory configuration	Investigated if the JA performance of CwASD is dependent on the social cues that the robot uses in the therapy sessions	Compared the behaviour of CwASD with that of TD children during a JA elicitation task	Improved social skills in CwASD using a long-term, in-home social robot	Designed a robot system that assisted in behavioral intervention programs of CwASD	Designed a fully autonomous closed-loop robotic system that can target core deficits of ASD
Robots and Control	Nao Wizard of Oz	Nao Wizard of Oz	Nao Wizard of Oz	CommU Wizard of Oz	Jibo Automated	iRobiQ and CARO Semi-automated	Nao Automated
Sensing system	Single RGB-D sensor	Two RGB-D sensors Three RGB cameras	Two RGB-D sensors Three RGB cameras	Single RGB camera	Two RGB cameras	Single RGB-D sensor	Four RGB cameras
Focus of Attention	Two images	Two objects on the table	N/A	Two images that were replaced in each session.	N/A	N/A	Two flat TVs
Prompts	Looking Looking + pointing Looking + pointing + talking	Looking + pointing	Looking Looking + pointing Looking + pointing + talking	Looking	Six prompts examine different aspects of JA, including gaze following, response to name and greeting	Discrete trial teaching: three term contingency antecedent stimulus (Sa), acceptable response (Ra), and consequent stimulus (Sc)	Looking + talking Looking + talking + pointing

Table 1 – Review of related works on joint attention assessment using robot-assisted intervention

	(ANZALONE *et al.*, 2018)	(CAI *et al.*, 2018)	(DAVID *et al.*, 2018)	(KUMAZAKI *et al.*, 2018a)	(SCASSELLATI *et al.*, 2018)	(YUN *et al.*, 2017)	(ZHENG *et al.*, 2018)
Metrics and Data Analysis	Response events, displacement, gazing and energy. Data were recorded and analyzed offline using manually video coding	JA, imitation and turn taking score. Automated estimation of metrics	Performance on the JA. Data were recorded and analyzed offline using manually video coding	Performance on the JA. Data were recorded and analyzed offline using manually video coding	Performance on the JA. Data were recorded and analyzed offline using manually video coding	Eye contact and facial emotion recognition. Data were recorded and analyzed offline using manually video coding	Performance on the JA and preferential attention. Data were recorded and analyzed offline using manually video coding
Children	ASD=25; 7.94 ± 1.67 TD = 12; 8.06 ± 2.49 ASD = 8 (6-month follow-up)	ASD = 5; 3-6 years	ASD= 5; 3-6 years	ASD=30; 5-6 years TD=38; 5-6 years	ASD= 12; 9.02 ± 1.41	ASD= 15 split into two groups: Therapy Group (n=8, 5.75 ± 0.89) and Control Group (n=7; 6.32 ± 1.23)	ASD= 14; 2.78 ± 0.65
Research Design	Single-subject design with single session In addition, they compared baseline and after six month of multiple session (four time a week at home and one time a week at hospital)	Single-subject design with single session	Single-subject design with Baseline Measurements (BM), Robot-Enhanced Treatment (RET) and Standard Human Treatment (SHT), eight sessions for each phase	Group-subject design with pretest, and post-test phases for each group of robotic intervention group (ASD =16 ;TD =17) Control group (ASD = 14; TD = 21)	Single-subject designs with pretest, test, and post-test phases, each phase lasting for about 30 days	Group-subject design with eight sessions	Single subject design with four sessions

Table 1 continued

	(ANZALONE *et al.*, 2018)	(CAI *et al.*, 2018)	(DAVID *et al.*, 2018)	(KUMAZAKI *et al.*, 2018a)	(SCASSELLATI *et al.*, 2018)	(YUN *et al.*, 2017)	(ZHENG *et al.*, 2018)
Findings	TD Children respond more than CwASD to the JA induction performed by the robot, in terms of head movements responses to JA induction. CwASD spend less time gazing towards the focus of attention than TD children and CwASD employ a higher amount of body energy than that of TD children. Features extracted during the JA task at six months from CwASD tend to change in a direction that is closer to TD value	There are no conclusions regarding the JA performance of children	The results for the first three participants shows that using more cues for prompting JA increases the performance of the children. Using pointing with or without vocal instructions, in addition to gaze orientation, leads to more JA engagement than gaze orientation alone. Robot interaction follows a similar pattern as human interaction, and pointing is a key ingredient to engage CwASD	JA in the CwASD was better during the robotic intervention than during the human agent intervention. These children exhibited improved performance in the JA task with human after interacting with the robot CommU. The facilitative effect of JA due to robot intervention was larger in CwASD than in TD children	CwASD showed improvement on JA skills with adults when not in the presence of the robot. Caregivers reported less prompting over time and overall increased communication. Caregivers reported improved social behaviour directed both toward themselves and toward others in areas including eye contact, initiation of communication, and responses to communication	After completing treatment, the eye contact percentages were significantly increased in both groups, but tended to gradually decrease to similar level as in the baseline in the follow-up. The treatment effects on the behavioral measures did not differ according to type of facilitator which suggested that robots are useful mediators of social skills training for CwASD	More instructive prompt levels on top of low prompt level elicited more target hits. The results showed that participants' JA skills improved significantly, throughout four sessions

Table 1 continued

Illustration by STEVE ASBELL; @rainforestgardn.
Award-winning author/illustrator who fights for inclusion in children's books.
Steve Asbell gave us the permission to use this illustration only for academic purposes

CHAPTER 3

MULTIMODAL ENVIRONMENT FOR ROBOT-MEDIATED INTERVENTION

This thesis aims to present a technological framework based on a Multimodal Environment for Robot-mediated Intervention (MERI) to assist clinicians with the ASD diagnostic process. The system was developed and implemented in two phases. In the first phase, technical requirements of low and middle-level, such as sensing, architecture, communication protocols, and algorithms for data processing, are implemented. In the second phase, user-friendly requirements are taken into accounts, such as user interfaces and advanced function for offline data processing and analysis.

In the first phase, the framework is composed of a responsive robotic platform, a flexible and scalable vision sensor network, and an automated face analysis algorithm based on machine learning models. In this phase, we take advantage of some neural models available as open sources projects to build an entirely new pipeline algorithm for global recognition and tracking of child's face under intervention among many faces present in a typical unstructured clinical scenario, to estimate the child's visual focus of attention along the time. In the second phase, a new approach of MERI is presented, which aims to wrapper into a single, and seamless user interface, the multiple stages of the autism diagnostic process, such as protocol planning and execution, data recording and analysis, as well as, the low and middle-level data processing. The main elements of the MERI system are presented as follow.

This chapter was adjusted from: Ramírez-Duque, A., Bastos, T.,Munera, M.,Cifuentes, C., Frizera-Neto, A., (2019). Robot-Assisted Intervention for Children with Special Needs: A Comparative Assessment for Autism Screening. *Manuscript submitted for publication*, and (RAMÍREZ-DUQUE *et al.*, 2019).

3.1 System Architecture Overview

The Robot Operating System (ROS) used in this work is a flexible and scalable open framework for writing modular robot-centered systems. Similar to a computing operating system, ROS manages the interface between robot hardware and software modules and provides common device drivers, data structures and tool-based packages, such as visualization and debugging tools. In addition, ROS uses an interface definition language (IDL) to describe the messages sent between process or nodes, this feature facilitates the multi-language (C++, Python and Lisp) development (QUIGLEY *et al.*, 2009).

The overall system developed here was built using a node graph architecture, taking advantages of the principal ROS design criteria. As with ROS, our system consists of a number of nodes to local video processing together a robot's behavior estimation, distributed around a number of different hosts and connected at runtime in a peer-to-peer topology. The inter-node connection is implemented as a hand-shaking and occurs in XML-RPC protocol along with a web-socket communication for robot's web-based node (/Ono_node, see Figure 1). The node structure is flexible, scalable and can be dynamically modified, i.e., each node can be started and left running along an experimental session or resumed and connected to each other at runtime.

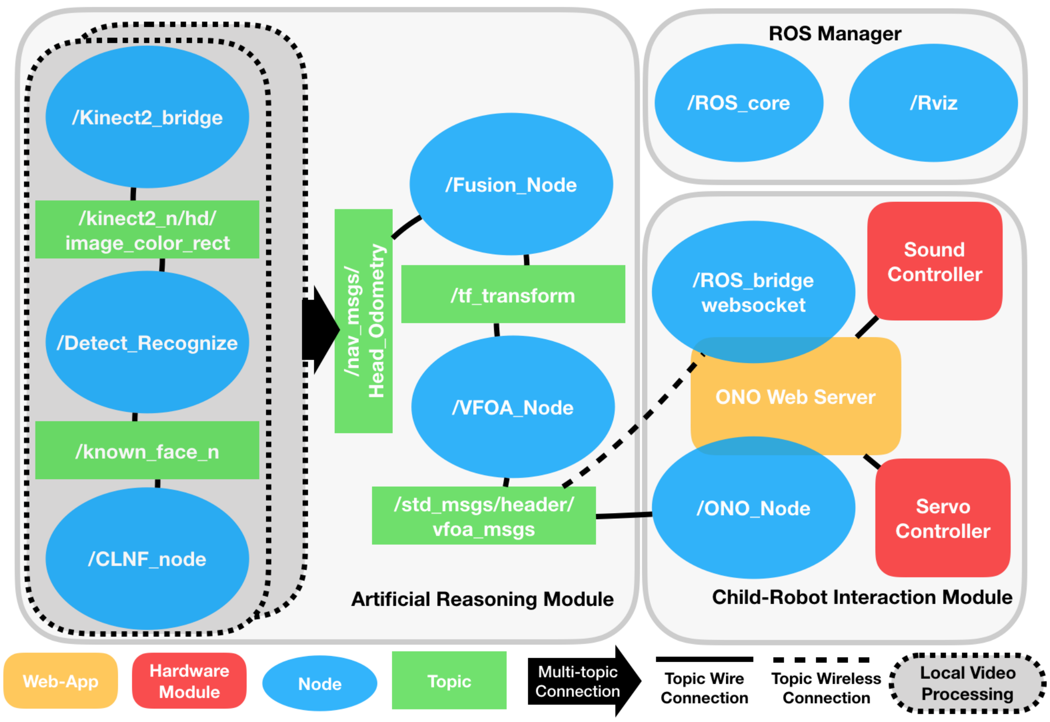

Figure 1 – Node graph architecture of the ROS-based MERI system. The system is composed of two interconnected modules, an artificial reasoning module and a CRI-channel module. The Ono web server has two way of bidirectional communication: a web-socket and a standard ROS Subscriber.

In addition, from a general perspective, any robotic platform with web-socket communication can be integrated. The developed system is composed of two interconnected modules as shown in Figure 1: an artificial reasoning module and a CRI-channel module. The module architectures are detailed in the following subsections.

3.1.1 Architecture of Reasoning Module

In this module, a distributed architecture for local video processing is implemented. The data of each RGBD sensor in the multi-camera system are processed for two nodes, in which the first is a driver level node and the second is a processing node. The driver[1] node transforms the streaming data of the RGBD sensor into the ROS messages format. The driver addresses the data through a specialized transport provided by plugings to publish images in a compressed representation while the receptor node only sees *sensor_msgs/Image* messages. The data processing node executes the face analysis algorithm. This node uses an *image_transport* subscriber and a ROS package called CvBridge to turn the data into an image format supported for the typical computer vision algorithms. Later, the same node publishes the head pose and eye gaze direction by means of a ROS navigation message defined as *nav_msgs/Odometry*.

An additional node hosted in the most powerful workstation carries out a data fusion of all navigation messages that were generated in the local processing stage. In addition to the fusion, this node computes the Visual Focus of Attention (VFOA) and publishes it as a *std_msgs/Header*, in which the time stamp and the target name of the VFOA estimation are registered. A schematic representation of the distributed system is shown in Figure 2

[1] Tools for using the Kinect One (Kinect V2) in ROS (https://github.com/code-iai/iai_kinect2).

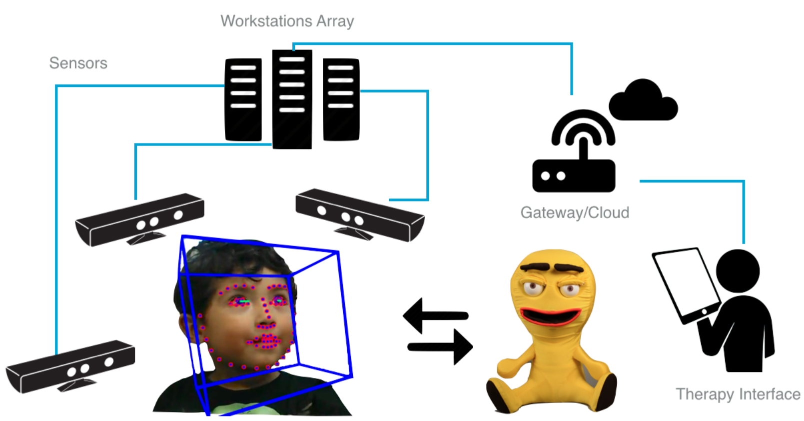

Figure 2 – Schematic representation of the distributed architecture for local video processing. The systems is composed of a RGBD sensor array, a workstation per sensor and a gateway for local interconnection.

3.1.2 *Architecture of CRI-Channel*

The system proposed here has two bidirectional communication channels, a robot-device, and a web-based application to interact with both the child and the therapist. The robot device can interact with the CwASD executing different physical actions, such as facial expression, upper limb poses, and verbal communication. Thus, according to the child's performance, the reasoning module can modify the robot's behavior through automatic gaze shifting, changing the facial expression and providing sound rewards. The client-side application was developed to allow the therapist to control and register all step of the intervention protocol. This interface was also used to supervise and control the robot's behavior and to offer feedback to the therapist about the child's performance along the intervention. This App has two channels of communication for interacting with the reasoning module. The first connection uses a web-socket protocol and a RosBridge_suite package to support the interpretation of ROS messages, as well as, JSON-based commands in ROS. The second one uses a ROS module developed in the server-side application to directly run a ROS node and communicate with standard ROS publishers and subscribers.

3.2 The Robotic Platform: Ono

The CRI is implemented through the Open Source Platform for Social Robotics (OPSORO)[2], which is a promising and straightforward system developed for face to face communication composed of a low-cost modular robot called Ono (see Fig. 3) and web-based applications (VANDEVELDE *et al.*, 2013). Some of the most important requirements and characteristics that make Ono interesting for this CRI strategy are explained in the following sections.

3.2.1 *Appearance and Identity*

The robot is covered in foam and also fabric to have a more inviting and huggable appearance to the children. The robot has an oversized head to make its facial expressions more prominent and to highlight the importance for communication and emotional interaction. As a consequence of its size and pose, children can interact with the robot at eye height when the robot is placed on a table.

The robot Ono has not a predefined identity, as the only element previously conceived

[2] Open Source Platform for Social Robotics (OPSORO) (http://www.opsoro.com).

Figure 3 – Ono robot, developed through the open source platform for social robotics (OPSORO).

is the name. Unlike other robots that have well-defined identities, such as Probo (VANDERBORGHT *et al.*, 2012) or Kaspar (DAUTENHAHN, 2007), in this dissertation the robot's identity and physical appearance is analyzed with the participation of the child through a co-creation process (See Chapter 6). For this reason, a neutral appearance is initially used. Beside, in the intervention, the therapist can provide the child with clothes and accessories to customize Ono.

3.2.2 *Mechanics Structure*

As the initial design of Ono is composed only of the actuated face, in this work it was needed to provide Ono with some body language. For this purpose, motorized arms were designed and implemented.

The new design of Ono has a fully face and two arms actuated, giving a total of 17 Degrees of Freedom (DOF). The Ono is able to perform facial expressions and nonverbal cues, such as waving, shake hands and pointing towards objects, moving its arms (2 DOF x 2), eyes (2 DOF x 2), eyelids (2 DOF x 2), eyebrows (1 DOF x 2), and mouth (3 DOF). The robot has also a sound module that allows explicit positive feedback as well as reinforcement learning through playing words, conversations and other sounds.

3.2.3 *Social expressiveness*

In order to improve social interaction with a child, the Ono is able to exhibit different facial expressions. The Ono's expressiveness is based on the Facial Action Coding System (FACS) developed in (EKMAN; FRIESEN, 1978). Each DOF that composes the Ono's face is linked with a set of Action Units (AU) defined by the FACS, and each facial expression is determined for specific AU values. The facial expressions are represented as a 2D vector $fe = (v, a)$ in the emotion circumplex model defined by valence and arousal (VANDERBORGHT *et al.*, 2012). In this context, the basic facial expressions are specified on a unit circle, where the neutral expression corresponds to the origin of the space $fe_0 = (0, 0)$. The relation between the DOF position and AU values is resolved through a lookup table algorithm using a predefined configuration file (VANDEVELDE *et al.*, 2013).

3.2.4 Adaptability and Reproducibility

The application of the Do-It-Yourself (DIY) concept is the principal feature of Ono's design, which facilitates its dissemination and use in research areas other than engineering as health care. These characteristics allow Ono building for any person without specialized engineering knowledge. Additionally, it is possible to replicate Ono without the need for high-end components or manufacturing machines (VANDEVELDE *et al.*, 2013). The electronic system is based on a Raspberry Pi single-board computer combined with a custom OPSORO module with circuitry to control up to 32 servos, drive speakers and touch sensors. Any sensor or actuator compatible with the embedded communication protocols (UART, I2C, SPI) implemented on the Raspberry Pi can be used by this platform.

3.2.5 Control and Autonomy

With the information delivered for the automated reasoning module, it was possible to automate the Ono's behavior and, then, the robot can infer and interpret the children's intentions to react most accurately to the action performed by them, thus enabling a more efficient and dynamic interaction with Ono. In this work, the automated Ono's behavior is partially implemented, i.e., the MERI system can modify some physical actions of Ono using the feedback information about the child's behavior. The actions suitable to be modified are gaze shift toward the child in specific events, changing from neutral to positive facial expression when the child looks toward the target, and providing sound rewards. Also, an Aliveness Behavior Module (ABM) is implemented to improve the CRI, which consist of blinking the robot's eyes and changing its arms among some predefined poses. Also, the robot can be manually operated through a remote controller hosted in the client-side application.

3.3 Reasoning Module: machine learning methods for child's face analysis

The automated child's face analysis consists of monitoring nonverbal cues, such as head and body movements, head pose, eye gaze, visual contact and visual focus of attention. In this work, a pipeline algorithm is implemented using machine learning neural models for face analysis. The chosen methods were developed using state-of-art trained neural models, available by Dlib[3] (KING, 2009) and OpenFace[4] (BALTRUŠAITIS *et al.*, 2016). Some modification such as turn the neural model an attribute of the ROS node class, and evaluate this in each topic callback, were needed to run the neural models into a common ROS node.

The algorithm proposed for child's face analysis involves face detection, recognition, segmentation and tracking, landmarks detection and tracking, head pose, eye gaze and VFOA estimation. In addition, the architecture proposed here also implement new methods for asynchronous matching and fusion of all local data, visual focus of attention estimation based on Hidden Markov Model (HMM) and direct connection with the CRI-channel to influence the robot's behaviors. A scheme of the pipeline algorithm is shown in Figure 4.

3.3.1 Child's face detection and recognition

The in-clinic setup requires differentiate the child's face from other faces detected and found in the scene. For this reason, a face recognition process was also implemented in this work. First, the face detection is executed to initialize the face recognition process and,

[3] Dlib C++ Library (http://dlib.net/).
[4] An Open Source Facial Behavior Analysis (https://github.com/TadasBaltrusaitis/OpenFace).

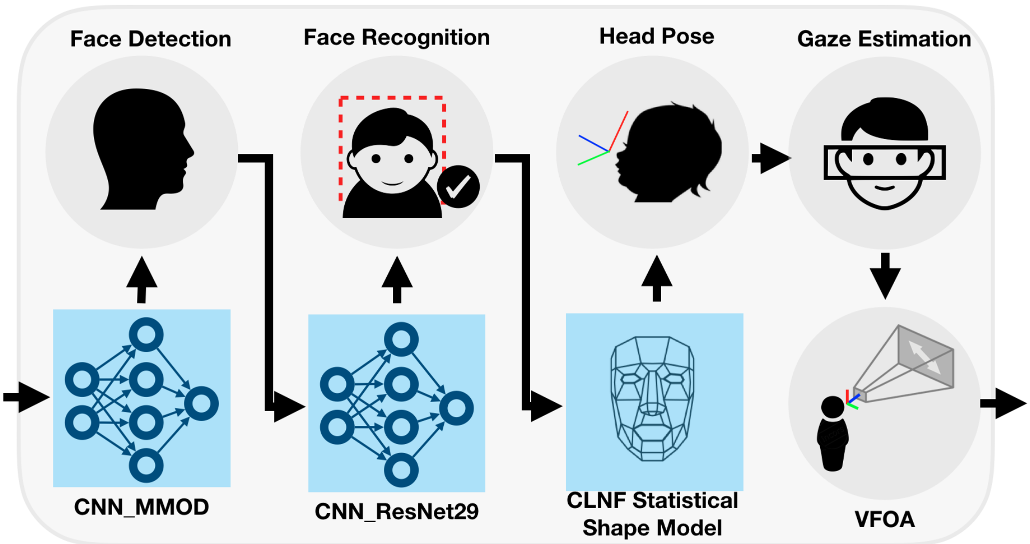

Figure 4 – Pipeline algorithm of the automated child's face analysis.

subsequently, initialize the landmarks detection. In this work, both detection and recognition are implemented using deep learning models, which are described in this section.

In the detection process, a Convolutional Neural Network (CNN) based face detector with a Max-Margin Object Detection (MMOD) as loss layer is used (KING, 2015). The CNN consists first of a block composed of three downsampling layers, which apply convolution with a 5x5 filter size and 2x2 stride to reduce the size of the image up to eight times its original size and generate a feature map with 16 dimensions. Later, the results are processed for one more block composed of four convolutional layers to get the final output of the network. The three first layers of the last block have 5x5 filter size and 1x1 stride, but, the last layer has only 1 channel and a 9x9 filter size. The values in the last channel are large when the network thinks it has found a face at a particular location. All convolutional block above are implemented with two additional layers among convolutional layers, pointwise linear transformation, and Rectified Linear Units (RELU) to apply the non-saturating activation function $f(x) = \max(0,x)$. The training dataset used to create the model is composed of 6975 faces and is available at Dlib's homepage[5].

The face recognition algorithm used in this work is inspired on the deep residual model from (HE *et al.*, 2016), where the authors reformulate the convolutional layers to learn a residual functions $F(x) := H(x) - x$ with reference to the layer inputs x, instead of learning unreferenced functions. In the practical implementation, the previous formulation means inserting shortcut connections, which turn the network into its counterpart residual version. The CNN model then transforms each face detected to a 128D vector space in which images from the same person will be close to each other, but faces from different people will be far apart. Finally, the faces are classified as child's face, caregiver's face and therapist's face.

Both detection and recognition CNN model were implemented and trained from (KING, 2009) and released in Dlib 19.6.

3.3.2 Face Analysis, Landmarks, Head Pose and Eye Gaze

This work uses the technique for landmarks detection, head pose and eye gaze estimation developed in (BALTRUŠAITIS *et al.*, 2013) and named Conditional Local Neural Fields (CLNF). This technique is an extension of the Constrained Local Model (CLM) algorithm using specialized local detectors or patch experts. CNLF model consists of a statistical shape model, which is learned from data examples and is parametrized for m components of linear

[5] Dlib's Data Set (http://dlib.net/files/data/dlib_face_detection_dataset-2016-09-30.tar.gz).

deformation to control the possible shape variations of the non-rigid objects (CRISTINACCE; COOTES, 2006). Approaches based on CLM (SARAGIH *et al.*, 2011; BALTRUŠAITIS *et al.*, 2012) and CLNF (BALTRUŠAITIS *et al.*, 2013) model the object appearance in a local fashion, i.e, each feature point has its own appearance model to describe the amount of misalignment.

CLNF-based landmark detection consists of three main parts: the shape model, the local detectors or patch experts, and the fitting algorithm, which are detailed below.

The CLNF technique uses a linear model to describe non-rigid deformations called Point Distribution Model (PDM). The PDM is used to estimate the likelihood of the shapes being in a specific class, given a set of feature points (CRISTINACCE; COOTES, 2006). This is important for model fitting and shape recognition.

The shape of a face that has n landmark points can be described as:

$$X = [X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_n, Z_1, Z_2, \ldots, Z_n], \tag{3.1}$$

and the class that describes a valid instance of a face using PDM can be represented as:

$$X = \bar{X} + \Phi q, \tag{3.2}$$

where $\bar{X}$ is the mean shape of the face, Φ describes the principal deformation modes of the shape, and q represents the non-rigid deformation parameters. Both $\bar{X}$ and Φ are learned automatically from labeled data using Principal Component Analysis (PCA). The probability density distribution of the instances into the shape class is expressed as a zero mean Gaussian with Covariance matrix $\Lambda = ([\lambda_1; \ldots; \lambda_m])$ evaluated at q:

$$p(\mathbf{q}) = \mathcal{N}(q; 0; \Lambda) = \frac{1}{\sqrt{(2\pi)^m |\Lambda|}} exp\left\{ -\frac{1}{2}(q^T \Lambda^{-1} q) \right\} \tag{3.3}$$

Once the model is defined, it is necessary to place the 3D PDM in an image space. The following equation is used to transform between 3D space to image space using weak perspective projection (SARAGIH *et al.*, 2011):

$$x_i = s \cdot R_{2D} \cdot (\bar{X}_i + \Phi_i q) + t, \tag{3.4}$$

where $\bar{X}_i = [\bar{x}_i, \bar{y}_i, \bar{z}_i]^T$ is the mean value of the i^{th} landmark. The instance of the face in an image is, therefore, controlled using the parameter vector $\mathbf{p} = [s, w, t, q]$, where q represents the local non-rigid deformation, s is a scaling term, w is the rotation term that controls the 2×3 matrix R_{2D}, and t is the translation term.

The global parameters are used to estimate the head pose in reference to the camera space using orthographic camera projection and solving the Perspective-n-Point (PnP) problem respect to the detected landmarks. The PDM used in (BALTRUŠAITIS *et al.*, 2016) was trained on two public datasets (BELHUMEUR *et al.*, 2013; LE *et al.*, 2012). It result in a model with 34 non-rigid (Principal modes) and 6 rigid shape parameters.

The patch experts scheme is the main novelty implemented in the CLNF model. The new Local Neural Field (LNF) patch expert takes advantage of the non linear relationship between pixel values and the patch response maps. The LNF captures two kinds of spatial characteristics between pixels, such as similarity and sparsity (BALTRUŠAITIS *et al.*, 2013).

LNF patch expert can be interpreted as a three layer perceptron with a sigmoid activation function followed by a weighted sum of the hidden layers. It is also similar to the first layer of a Convolutional Neural Network (BALTRUŠAITIS *et al.*, 2016). The new LNF patch expert is able to learn from multiple illuminations and retain accuracy. This becomes important when creating landmark detectors and trackers that are expected to work in unseen environments and on unseen people.

The learning and inference process is developed using a gradient-based optimization method to help in finding locally optimal model parameters faster and more accurately.

In the CLNF model implemented in (BALTRUŠAITIS *et al.*, 2016), twenty-eight set in total of LNF patch experts were trained for seven views and four scales. The framework uses patch experts specifically trained to recognize the eyelids, iris and the pupil, in order to estimate the eye gaze (BALTRUŠAITIS *et al.*, 2016).

For each new image or video frame, the fitting algorithm of CLNF-based landmark detection process attempts to find the value of the local and global deformable model parameters $\mathbf{p}$ that minimizes the following function (SARAGIH *et al.*, 2011):

$$\mathscr{E}(\mathbf{p}) = \mathscr{R}(\mathbf{p}) \sum_{i=1}^{n} \mathscr{D}_i(x_i; \mathscr{I}), \tag{3.5}$$

where $\mathscr{R}$ is a weight to penalize unlikely shapes, which depends on the shape model, and $\mathscr{D}$ represents the misalignment of the i^{th} landmark in the image $\mathscr{I}$, which is function of both parameters **p** and the patch experts. Under the probabilistic point of view, the solution of (3.5) is equivalent to maximize the *a posteriori* probability (MAP) of the deformable model parameters **p**:

$$p(\mathbf{p} \mid \{l_i = 1\}_{i=1}^{n}, \mathscr{I}) \propto p((p)) \prod_{i=1}^{n} p(l_i = 1 \mid x_i, \mathscr{I}), \tag{3.6}$$

where, $l_i \in \{1, -1\}$ is a discrete random variable indicating whether the i^{th} landmark is aligned or misaligned, $p(\mathbf{p})$ is the prior probability of the deformable parameters **p**, and $p(l_i = 1 \mid x_i, \mathscr{I})$ is the probability of a landmark being aligned at a particular pixel location x_i, which is quantified from the response maps created by patch. Therefore, the last term in (3.6) represents the joint probability of the patch expert response maps.

The MAP problem is solved using an optimization strategy designed specifically for CLNF fitting called non-uniform regularized landmark mean shift (NU-RLMS) (BALTRUŠAITIS *et al.*, 2013), which uses two step process. The first step evaluates each of the patch experts around the current landmark using a Gaussian Kernel Density Estimator (KDE). The second step iteratively updates the model parameters to maximize (3.6).

The NU-RLMS uses expectation maximization algorithm, where the E-step involves evaluating the posterior probability over the candidates, and the M-step finds the parameter updated through the mean shift vector **v**. The mean shift vector points in the direction where the feature point should go, but the motion is restricted by the statistical shape model and the $\mathscr{R}(\mathbf{p})$. This interpretation leads to the new update function:

$$\arg\min_{\Delta\mathbf{p}} \left\{ \|J\Delta\mathbf{p} - \mathbf{v}\|_W^2 + r\,\|\mathbf{p} + \Delta\mathbf{p}\|_{\tilde{\Lambda}^{-1}}^2 \right\}, \tag{3.7}$$

where r is a regularization term, J is the Jacobian, which describe how the landmarks location are changing based on the infinitesimal changes of the parameters **p**, $\tilde{\Lambda}^{-1} = diag([0;0;0;0;0;0;\lambda_1^{-1};\cdots;\lambda_m^{-1}])$, and W allows for weighting of mean-shift vectors. Non-linear least squares leads to the following update rule:

$$\Delta\mathbf{p} = -\left(J^T W J + r\Lambda^{-1}\right)\left(r\Lambda^{-1}\mathbf{p} - J^T W \mathbf{v}\right). \tag{3.8}$$

To construct W, the performance of patch experts on training data is used.

3.3.3 Data Fusion

The fusion of the local results for the head pose estimation is done applying a consensus over the rotation algorithm (JORSTAD *et al.*, 2010). This algorithm consists of calculating the weighted average pose between each camera estimation and its immediate sensors' estimation neighbors using the axis-angle representation.

Given the angle θ_i and the normalized axis of rotation $\vec{u}_i$ of the head pose in the camera i then the average global pose can be expressed as:

$$\vec{u}_{\text{sum}} = \sum_{i=1}^{N} \omega_c \omega_d \left(\theta_i \vec{u}_i\right) \tag{3.9}$$

$$\theta_{\text{ave}} = \frac{1}{N} \|\vec{u}_{\text{sum}}\| \tag{3.10}$$

$$\vec{u}_{\text{ave}} = \frac{1}{\theta_{\text{ave}} N} \vec{u}_{\text{sum}}\,; \tag{3.11}$$

where N is the number of sensor with successful local pose estimation.

Each local pose is penalized by two weights, ω_c and ω_d which represent the alignment confidence of landmarks detection procedure and the Mahalanobis distances between the head pose and a neutral pose.

3.3.4 Field of View and Visual focus of Attention

The VFOA estimation model is implemented as a dynamic Bayesian network through a HMM. The model assumes a specific set of child's attention attractors or targets $\mathbb{F}$. The estimation process decodes the sequence of child's head poses $H_t = (H_t^{yaw}, H_t^{pitch}) \in \mathbb{R}^2$ in terms of VFOA states $F_t \in \mathbb{F}$ at time t (BA; ODOBEZ, 2008). The probability distribution of the head poses in reference to a given VFOA target is represented by a Gaussian distribution, whereas the transitions among these targets are represented by the transition matrix A. The HMM equations can then be written as follows:

$$P(H_t \mid F_t = f, \mu_t^h) = \mathcal{N}(H_t \mid \mu_t^h(f), \Sigma_H(f)) \tag{3.12}$$

$$p(F_t = f \mid F_{t-1} = \hat{f}) = A_{f\hat{f}}. \tag{3.13}$$

The Gaussian covariances are defined manually to reflect target sizes and head pose estimation variability. Moreover, the Gaussian means corresponding to each specific target μ_t^h

is calculated through a gaze model that sets this parameter as a fixed linear combination of the target direction and the head reference direction (SHEIKHI; ODOBEZ, 2015):

$$\mu_t^h(f) = \alpha \star \mu_t(f) + (1_2 - \alpha) \star R_t, \tag{3.14}$$

where $\star$ denotes the component wise product $1_2 = (1,1)$, $\alpha = (\alpha^{yaw}, \alpha^{pitch}) = (0.7, 0.5)$ are adjustable constants that describe the fraction of the gaze shift that corresponds to the child's head rotation, $\mu_t \in (\mathbb{R}^2)^K$ is the directions of the given K targets, and $R_t \in \mathbb{R}^2$ represents the reference direction, which is the average head pose over a time window W^R.

The above assumption describes the body orientation behavior of any child who tends to orient himself/herself towards the set of gaze targets to make more comfortable to rotate his/her head towards different targets (SHEIKHI; ODOBEZ, 2015).

$$R_t = \frac{1}{W^R} \sum_{i=t-W^R}^{t} H_i. \tag{3.15}$$

Finally, for the estimation of the VFOA sequence a classic Viterbi algorithm of HMM is implemented (BA; ODOBEZ, 2008).

3.4 Phase two of the MERI system

The current MERI version remains as a ROS-based[6] system, but now with a user-friendly approach. Thus, in this phase all the execution processes and operation commands were assembled in a common framework using qt-based GUI development for ROS[7]. The ROS metapackage, named rqt, provides a GUI that enables multiple qt widgets with ROS capabilities to be docked in a single window. The main elements developed in the second phase are described as follows:

3.4.1 MERI GUI Interface

The MERI GUI was developed to be compatible and portable like all rqt tools of the rqt_common_plugins[8] packages, which means that this interface can be used in conjunction with

6 Robot Operating System, http://www.ros.org
7 ROS metapackage for rqt interface, http://wiki.ros.org/rqt
8 RQT common plugins, http://wiki.ros.org/rqt_common_plugins

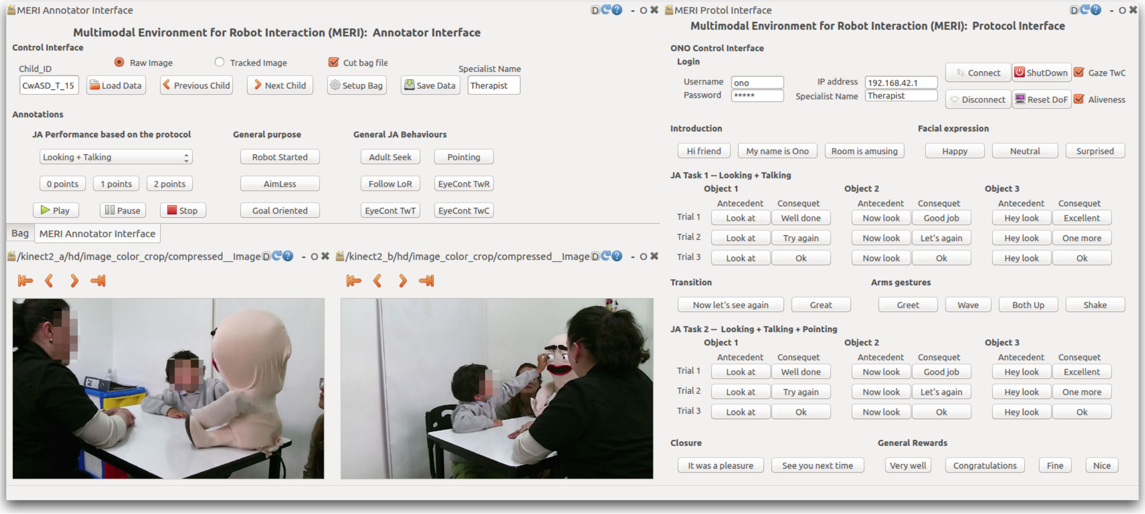

Figure 5 – MERI GUI Interface used in conjunction with the visualization plugin.

any of the existing plugins. The MERI GUI is composed of two widgets designed to provide the therapist or engineer staff with a tool capable of simultaneously monitoring the session, record the sensor data, operate the robot, and make online annotations about the behavior of the participant throughout the intervention.

The MERI Annotator Interface (MERI-AI) and the MERI Protocol Interface (MERI-PI) are the two developed widgets that are first presented in this work and are shown in Figure 5. MERI-PI was developed to control the execution of the clinical protocol and adjust the behavior of the robot in a supervised approach. This widget uses a web-application architecture that allows connecting to any robotic platform with a web server side implementation for robot controller.

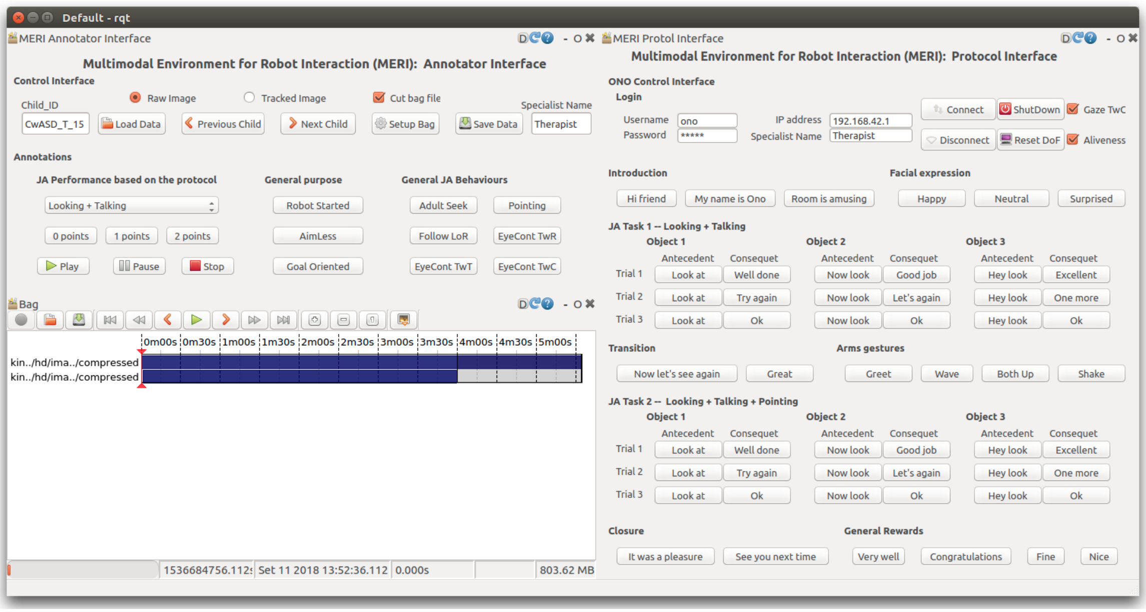

Figure 6 – MERI Interface, which is composed of MERI annotator interface (MERI-AI) and MERI protocol interface (MERI-PI).

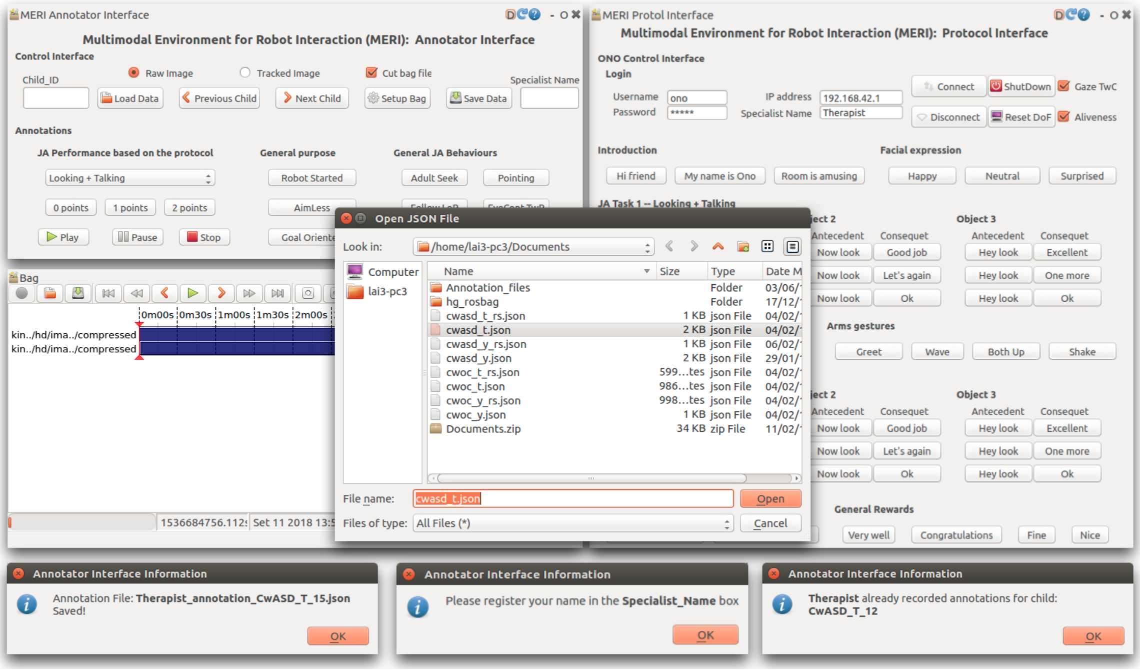

Figure 7 – MERI GUI Interface tool used for load and save the bag files and the annotations files, as well as, some dialog boxes designed to assist the user.

Therefore, the widget has fields for the IP address as well as log-in information to access all controller functions. The widget's core is composed mainly of a route map explicitly developed to execute JA tasks, i.e., a sequence of buttons that controls semi-automatically and in sequential order all the actions performed by the robot to elicit JA behaviors. Other commands to control the activation of automatic behaviors, such as greeting by waving hands, saying goodbye with a handshake, making facial expressions and activating aliveness mode were also added within the widget.

MERI-AI was developed to bring together several technical demands from therapists and clinical staff to facilitate both online and offline analysis of the sessions. The main task of this interface is to provide the therapist with the ability to make notes about the behavior of children with a single click. In addition, the therapist can directly control the functions of video recording the session in a bag file, which is a file format for storing the ROS messages. Then, the therapist also can execute actions such as play, pause, or resume the bag file recorded (see Figure 6). To facilitate the task of video coding the annotations are split into three blocks: (i) specific annotations to describe the behavior and child's performance along of JA task; (ii) annotations to quantify typical variables, such as eye-contact and focus of attention; and (iii) general purpose annotations found in the specific literature to describe CRI scenarios (LEMAIGNAN *et al.*, 2017).

The MERI system records all the annotations generated by the therapist throughout the session and stores these in a JSON file. The actions performed by the robot/therapist in the protocol interface are also stored in a JSON file. Each annotation or action data is composed of an identifier field, a description tag, and the ROS timestamp related to the instant that was generated. This approach allows achieving a seamless data structure to analyze without synchronization failures all multimedia data recorded (audio, video, annotations and logs actions). The GUI tool used for load and save the bag files and the annotations files, as well as, some dialog boxes are shown in Figure 7.

3.4.2 Sensing system and data processing

The sensing system consists of a multi-camera array, which acquires specific children's patterns. These patterns are related to the motor processes (point, head turn, gaze shift, embodied movements), visual-motor integration (a point at what peers are looking at) and aspects of visual attention (select the target object in the visual field).

The video data pipeline is composed of face localization, face template-based recognition, face tracking, landmarks detection and tracking, head pose and visual focus of attention estimation. A scheme of the video data processing is shown in Figure 8. The face processing algorithms uses the open machine learning frameworks developed in (BALTRUŠAITIS *et al.*, 2016; KING, 2009). All algorithms were implemented with support for GPU acceleration and can be run on local GPU boards or Cloud GPU containers. The main improvements to the sensing system are described below.

In the first version of MERI, before each session, an image of the participant's face is needed to be captured with an external camera and then downloaded to the main computer and distributed among the workstation manually. This procedure was tedious and was considered one of the weaknesses of the system. Therefore, the new pipeline implementation allows capturing automatically, in the first seconds of the interaction, an image of the child's face to be used as a seed for the creation of a template necessary in the recognition phase.

In the face recognition phase, instead of using a single image as in our previous version, the recognition process used as input of the CNN_ResNet29[9] model a face template formed by multiple images. The template is built using references images or keyframes of the child's face in different head orientation. Over the time of processing, the algorithm maintains a

[9] Dlib c++ Library, http://dlib.net

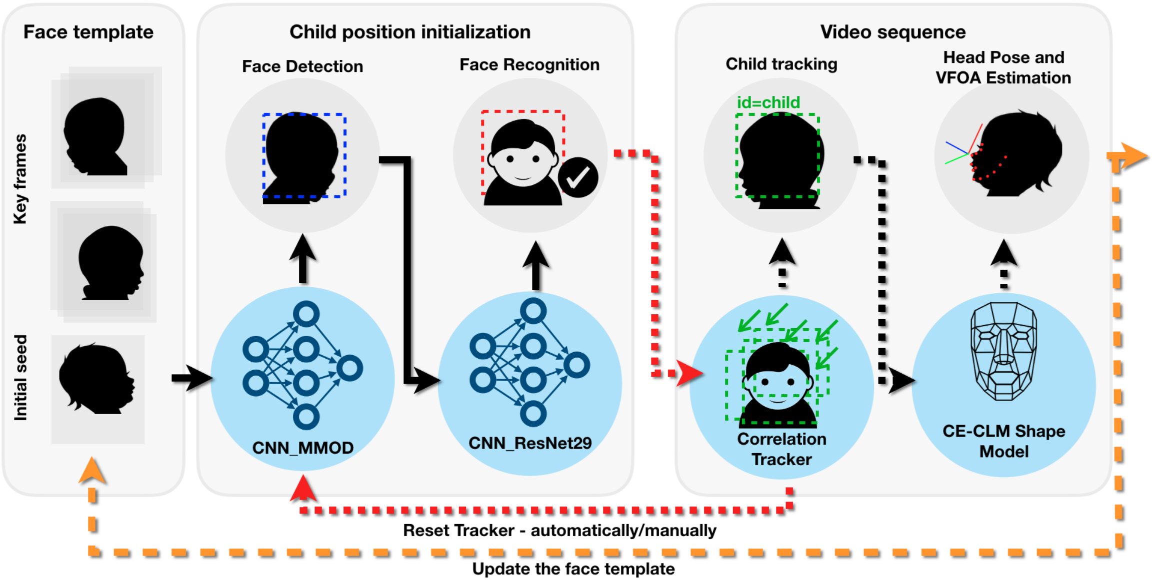

Figure 8 – Video data processing scheme of MERI system.

set of ten keyframes extracted from the different camera and time instants with a variety of head poses. Each new keyframe is stored if it contains a different orientation and high confidence. This improvement increases the recognition rate in self-occlusion scenarios.

After recognition stages, a child's face tracking process was used. The algorithm is based on the correlation tracker from the Dlib C++ library, which is an implementation of the method described in (DANELLJAN *et al.*, 2014). Face tracking requires execution of face recognition process only in the initialization of the pipeline. Performance of the tracking is considerably faster than face recognition, but less accurate. For this reason, it was necessary to implement an interruption to reinitialize the tracking by rerunning the face recognition. The reset by interruption is activated automatically when the size of the face bounding box increased by more than 15 % compared with the first estimation, or manually through a keyboard input by the operator.

Following the tracking stages, the face landmarks are detected and then the head orientation and the gaze vector are estimated. For these last phases, the new OpenFace 2.0 open framework version was used (BALTRUŠAITIS *et al.*, 2013), which instead of using the CLNF model (BALTRUŠAITIS *et al.*, 2013), uses a more modern version called Experts Constrained Local Model (CE-CLM) developed in (ZADEH *et al.*, 2017) and implemented in (BALTRUSAITIS *et al.*, 2018). This last version of OpenFace brings improvements regarding real-time performance and detection of faces in-the-wild. Finally, the head orientation and gaze estimation are correlated with the contextual information of the therapy scenario, and the current focus of attention is calculated.

3.4.3 Input-Output Visual focus of Attention Estimation

The current module for VFOA estimation leverages seamless feature of the MERI interface, which allows the system to be fully aware of therapist/robotic actions as well as the child's reactions. This provides more conveniently interpretation of the non-verbal cues performed by interacting children. Thus, the VFOA estimation model is implemented as an Input-Output Hidden Markov Model (IO-HMM), which uses contextual observation variables related to the mediator actions in order to estimate the VFOA. Thus, the robot conversation and actions context C_t appears as an input observation and provides expectations about which VFOA should be expected (SHEIKHI; ODOBEZ, 2015).

Thus, the probability distribution of the head poses in reference to a given VFOA target is represented as:

$$p\left(F_{1:t}|H_{1:t},C_{1:t},\mu^h_{1:t},R_{1:t}\right) \propto \prod_{t=1:t} p\left(H_t|F_t,\mu^h_t\right) p\left(F_t|F_{t-1},C_t\right), \tag{3.16}$$

and the Equations 2.9 and 2.10 are rewritten using the IO-HMM model as follows:

$$p(H_t \mid F_t = f, \mu^h_t) = \mathcal{N}(H_t \mid \mu^h_t(f), \Sigma_H(f)) \tag{3.17}$$

$$p\left(F_t \mid F_{t-1},C_t\right) \propto p(F_t = f \mid F_{t-1} = \hat{f})p\left(F_t = f \mid C_t = c\right) = A_{f\hat{f}}B_{cf} \tag{3.18}$$

In the robot-assisted diagnostic scenario, the robot interaction contexts are defined as temporal segments of attention elicitation toward a specific target called the antecedent stimulus (Sa). In this setup, the contextual events are automatically derived from the MERI-PI and the robot internal system, meaning that probabilities of the expected VFOA in time t can be modified in terms of previous robot's actions described in the matrix B_{cf}, which is defined depending on behavioral child's features. In this work, a learning approach using manual coding of ten previous interaction of TD children was used to determine the aforementioned parameter. The Gaussian covariances and the Gaussian mean are defined in the same way of Section 2.3.4.

Finally, using the outputs of the IO-HMM module, the last stage of the VFOA algorithm is to register the time and frequencies of each event of: looking toward targets and looking toward the mediators, in order to estimate JA related metrics, such as for JA score, eye contact, and adult seeking, which are described in Chapter 4.

CHAPTER 4

THE MERI SYSTEM IN-CLINIC SETUP ASSESSMENT

Throughout the development of this research, two activities of testing and validation took place. The first validation activity consisted of a case study conducted in Vila Velha-ES, Brazil. This case study aimed to validate the technological tools developed in the first two years, analyzing the methods feasibility as well as the designed clinical protocol. The case study was conducted at the Clinic of the University of Vila Velha (UVV) between September and December 2017. Once the case study was completed, all aspects, both clinical and technical, were assessed. Thus, in the second phase of this research, a redesigned clinical intervention with a higher number of participants was implemented. Given the difficulty of convening a significant amount of children for the new validation stage, it was decided to contact the Howard Gardner (HG) clinic in Colombia. Thus, between April and September 2018, this research was conducted in Colombia at the headquarters of the HG clinic and with the support of the Colombian School of Engineering Julio Garavito. The experimental and clinical details, as well as the main findings, are described in this chapter.

4.1 Robot-Assisted Intervention: A Case Study

A multidisciplinary team of psychologists, doctors and engineers developed a case study using a psychology room at the UVV clinic, equipped with a unidirectional mirror to

This chapter was adjusted from: Ramírez-Duque, A., Bastos, T.,Munera, M.,Cifuentes, C., Frizera-Neto, A., (2019). Robot-Assisted Intervention for Children with Special Needs: A Comparative Assessment for Autism Screening. *Manuscript submitted for publication*, and (RAMÍREZ-DUQUE *et al.*, 2019).

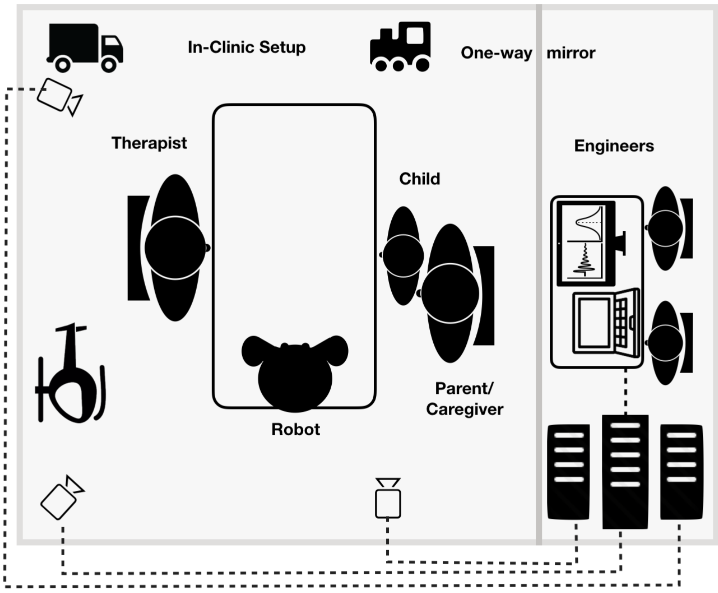

Figure 9 – Representation of the interventions room of in-clinic setup. The room was equipped with a table where the robot was placed and three toys were attached to room's walls.

perform behavioral observation appropriately. The room was prepared with a table and three chairs: one for the child, another for the caregiver and a third one for the therapist. The robot was placed on the table, and the following toys, a helicopter, a truck and a train, were attached to room's walls. The RGBD sensors were located close to the walls, and no additional camera was placed on the robot or the table, so as not to attract the child's attention. A representation of the interventions room of in-clinic setup is shown in Figure 9.

The vision system was composed of three Kinect V2 sensors. Each sensor was connected to a workstation equipped with a processor of Intel Core i5 family and a GeForce GTX GPU board (two workstations with GTX960 board, and one workstation with GTX580 board). All workstations were connected through a local area network, synchronized using the NTP protocol[1]. The sensors were intrinsically and extrinsically calibrated through a conventional calibration process using a standard black-white chessboard[2].

4.1.1 Intervention protocol

For the scope of case study, a specific clinical setup intervention to assess joint attention behaviors is presented. The therapist guides the intervention all the time and leverages the robot device as an alternative channel of communication with the child, thus, both the specialist and the robot remained in the room during the intervention. The children were

1 Network Time Protocol Homepage, (http://www.ntp.org).
2 Tools for using the Kinect One (Kinect V2) in ROS (https://github.com/code-iai/iai_kinect2).

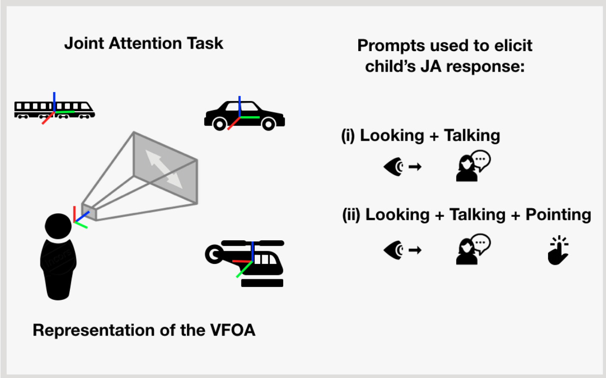

Figure 10 – Graphical representation of the JA task and the prompts used to elicit child's JA behaviors.

accompanied throughout the session by a caregiver who was oriented not to help the child in the execution of the tasks. The exercise developed aimed to direct the attention of the child towards objects located in the room through stimuli, such as, look at, point and speak. The stimuli were generated first only by the therapist and later just by the robot. A graphical representation of the JA task and the prompts used to elicit child's JA behaviors is shown in Figure 10.

4.1.2 Participants

Ten children without confirmed ASD diagnosis, but with evidence of risk factors, and three typically developing children as the control group participated in the experiments. All volunteers participated with their parent's consent, which were eleven boys (9 ASD, 2 TD) and two girls (1 ASD, 1 TD), between 36 months to 48 months. Each volunteer participated in one single session. The goal was to analyze the based-line of the child's behavior and establish differences in the behavioral reaction between TD and ASD children for stimuli generated through CRI and leverage the novelty effect raised by the robot mediator.

4.2 Results of case study

The child's nonverbal cues elicited by the CRI can be observed in Figure 11. Some examples of children's behavior tagged to perform the behavioral coding are shown in the six pictures. The tagged behaviors were: to look towards an object, towards the robot, and towards the therapist, to point and, to respond to a prompt of both mediators and self occlusion. Typical occlusion problem, as occlusion by hair, hands and the robot were detected.

The performance of video processing in the proof of concept session is reported in Figure 12, in which the results of the child's face detection and recognition, landmarks detection,

Figure 11 – The child's nonverbal cues elicited by the CRI, to look towards the therapist, towards the robot, point and self occlusion.

head pose and eye gaze estimation from different viewpoints are exhibited. The recognition process was able to detect all faces in the session successfully in most cases.

The child's head pose was captured throughout the session and analyzed automatically to estimate the evolution over time of child's head and the VFOA. Along the session, the child's neck right/left rotation movement was predominant (Yaw axis), while the neck flexion/extension (Pitch axis) and neck R/L lateral flexion movements (Roll axis) remained approximately constant. The Yaw rotation (angle amplitude grows in the robot's direction) of the TD children group is reported in Figure 13. The vertical light blue stripe indicates the intervention period with therapist-mediator, and the vertical light green stripe represents the period with robot-mediator. The continuous blue line represents the raw data recorded, and the continuous red line describes the average data trend. From the observation of the three plot, the TD children started the

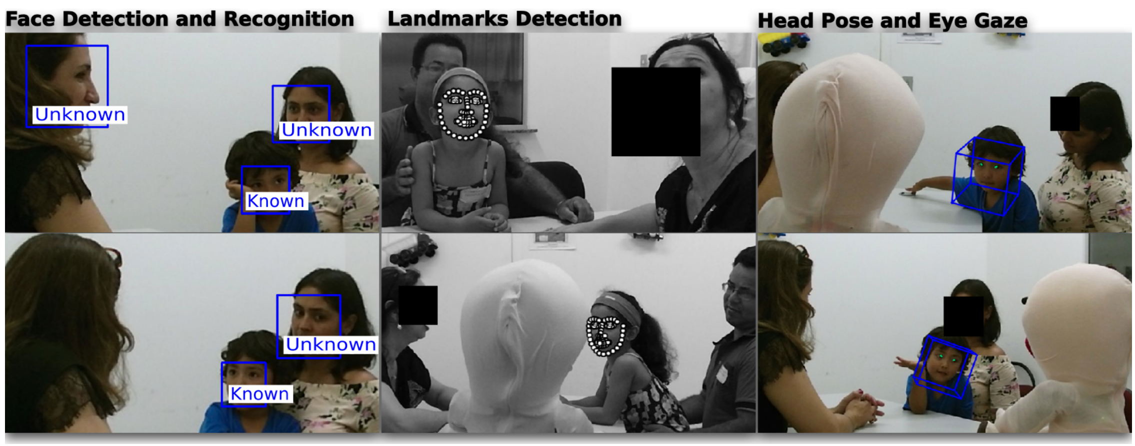

Figure 12 – Performance of the child's face analysis pipeline for the case study. Face detection and recognition, landmarks detection, head pose and eye gaze estimation were executed.

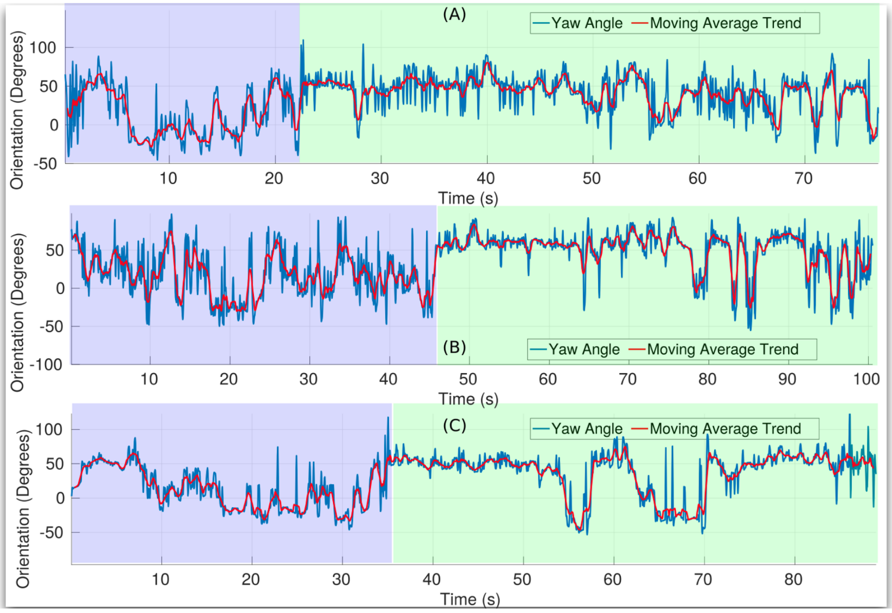

Figure 13 – Evolution over time of the child's head/neck rotation (yaw rotation) for a TD group. In light blue the therapist interval and in light green the robot mediated interval.

intervention looking towards the robot, evidently, the robot was a naturalistic attention attractor. Subsequently, when the therapist begins the protocol explaining the tasks, the children attention shifts towards the therapist. The children remained this behavior until that therapist introduced the robot-mediator. In this transition, the children's behaviors, such as, initiation of JA and Responding to a bid for JA were observed. Once the therapist changed the mediation with the robot, the children turned his/her attention to the robot and the objects in the room.

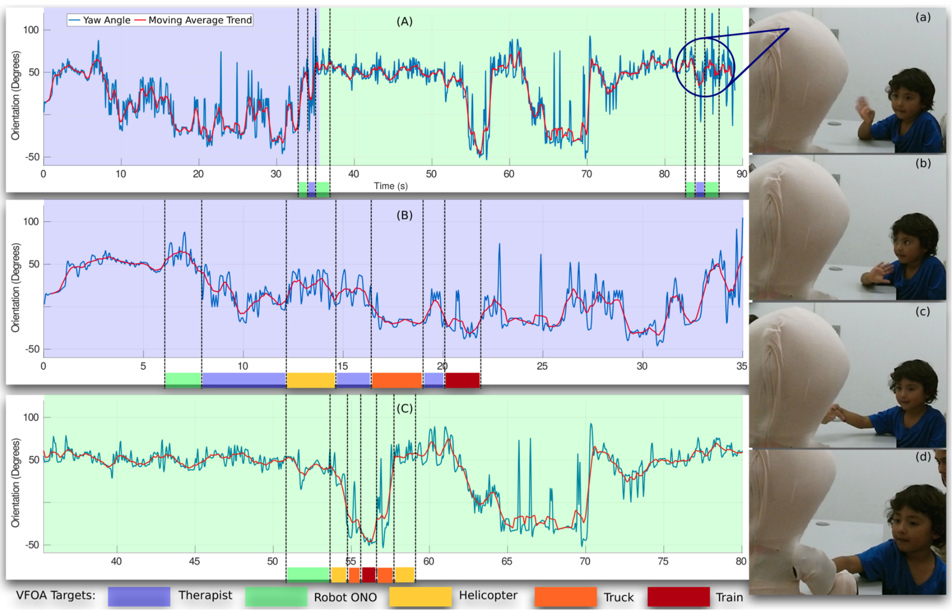

Figure 14 – Evolution over time of the child's head/neck rotation (yaw rotation) for a TD volunteer and VFOA estimation results. In light blue the therapist interval and in light green the robot mediated interval.

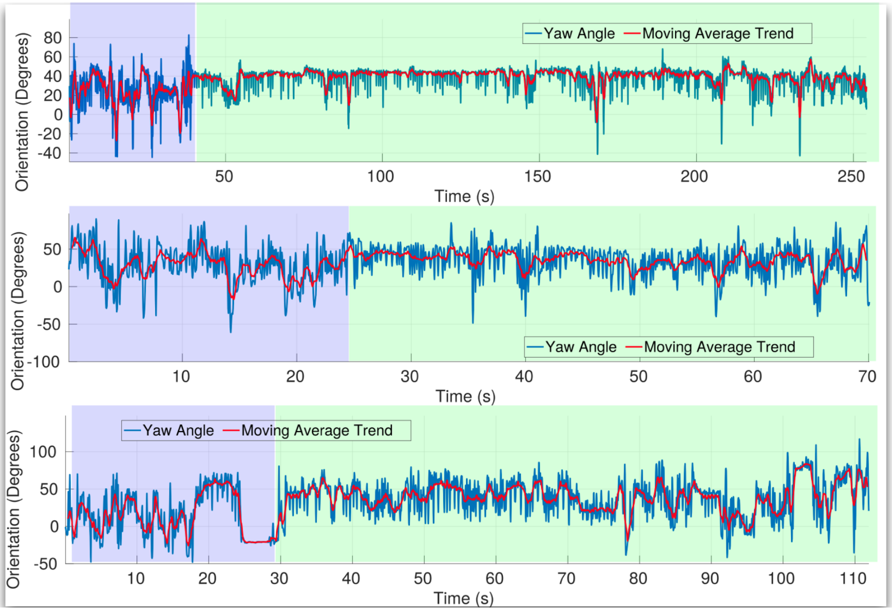

Figure 15 – Evolution over time of the child's head/neck rotation (yaw rotation) for a ASD group. In light blue the therapist interval and in light green the robot mediated interval.

A more detailed analysis of one of the TD volunteers is shown in Figure 14. The plot (A) shows the overall intervention session; the plot (B) and plot (C) are a zoom of the period with therapist and robot mediator, respectively. The colors convention in the three plots of Figure 14 describes the results generated by the automated estimation of VFOA. From these scenarios, some essential aspects already emerge. In the therapist-mediator interval, child responded to JA task using only one repetition for all prompt level. Child's behavior of RJA was according to the protocol, i.e., the child looked towards the therapist to wait for instructions, rapidly child searched in the target, and next looked again toward the therapist (Color sequence: light blue - yellow - light blue - orange - light blue - red). This behavior was the same for all prompts. In contrast, with the robot-mediator, the child did not look toward the robot among indications at consecutive targets (Color sequence: light green - yellow - orange - red - orange - yellow). The above happened because, in the protocol, both mediators executed the instructions in the same order, and the child memorized the commands and the object's position until the robot mediator interval. This fact did not affect the intervention's aim, as the robot mediator succeeded to elicit the child's behaviors of RJA and IJA. In addition, as highlighted in the plot (A) in Figure 14, when the session finalized and the robot mediator said goodbye, again, RJA and IJA behaviors were perceived. The pictures (a-d) show these events: first the child said goodbye towards the robot, then, he looked the therapist to confirm that the session ended and looked again towards the robot, finally the child took the robot's hand.

From the analysis of the three TD volunteers, the same reported behaviors were perceived. However, the analysis of the children in the ASD group showed different behavior patterns concerning comfort, visual contact and novelty stimulus effect during the sessions. The evolution over time of the child's head/neck rotation (yaw rotation) for an ASD group is shown in Figure 15. On the one hand, the ten children in the ASD group maintained more visual contact with the robot compared to the therapist and exhibited more interest in the robot platform compared to the TD children. However, the performance of the children in the activities of JA did not improve significantly when the robot executed the prompt. On the other hand, the clinicians manifested that in all cases the first visual contact toward them occurred in the instant that the robot entered the scene and started interacting, i.e., the Ono mediation elicited behaviors of IJA towards the therapist. In addition, the CwASD exhibited less discomfort regarding the session, from the first moment when the robot initiated mediation in the room and, in some cases, when showed appearance of verbal and non-verbal pro-social behaviors. These facts did not arise with the TD children, because the first visual contact with the therapist occurred when they entered the room. Additionally, TD children showed the ability to divide the attention between the robot and the therapist from the beginning to the end of the intervention, exhibiting comfort in every moment. The behavior modulation of CwASD is observed in Figure 15. Before the period with robot-mediator the children exhibited discomfort (unstable movements of their head), and after of this period, the head movement tended to be more stable.

The novelty of a robot-mediator at diagnostic session can be analyzed as an additional stimulus of the CRI. Accordingly, in this case study the children of the ASD group showed more behavior modification (attention and comfort) produced by the robot interaction at the beginning of the CRI, remaining until the end of the session. On the other hand, the children of the TD group responded to the novelty effect of the robot mediator from the time the child entered the room and saw the robot, until the beginning of the therapist presentation. For the above, despite the novelty of the stimuli effect, these did not seem to affect the social interaction between the TD children and the therapist, and in contrast, these stimuli seemed to enhance the CwASD social interaction with the therapist along the intervention.

These results are impressive, since they show the potential of CRI intervention to systematically elicit differences between the pattern of behavior on TD and ASD children. In the session, RJA and IJA were identified toward the therapist at the beginning of the intervention, at the transition between therapist to robot mediator, and at the end for all TD children. In contrast, IJA towards the therapist only was identified in the transition between mediators, for

ASD children. This fact shows a clear difference of behavior pattern between CwASD and TD children, which can be analyzed using a JA task protocol. In fact, these pattern differences can be used as evidence to improve the ASD diagnosis.

4.3 Robot-Assisted Intervention: Risk Factor Identification and Autism Screening

The MERI system developed in the second phase was adapted and installed in the music-therapy room of the Colombian Rehabilitation Clinic, Howard Gardner. The HG clinic specializes in therapies for children with special needs and with neurodiversity conditions. In this therapeutic institute, four kinds of personalized interventions are supplied: (i) occupational therapy, (ii) speech and language therapy, (iii) physical therapy, and (iv) psychology, which are given daily to each of their patients (young individuals from 2 to 18 years old). All procedures developed in this study were approved by the Colombian School of Engineering Julio Garavito Ethical Board.

4.3.1 Participants

Two groups of children were included in the study. The first group was formed by children (a) diagnosed with ASD (including pervasive developmental disorder not otherwise specified, Asperger's syndrome, autism) by an expert clinician based on the DSM-V criteria; and (b) without severe auditory or visual impairments. The Control Group (CG) was conformed by children without ASD, but with one of the following conditions: down syndrome, congenital hydrocephalus, or general learning disability. The inclusion criteria for the CG were: (a) diagnosed without ASD based on the DSM-V criteria, and (b) without severe auditory or visual impairments. Besides, HG specialists made a special selection of the children in CG to assure that they preserve good social skills. In this study, written informed consent was obtained from all of the parents or children who were able to participate.

Initially, 29 CwASD (22 boys and 7 girls; 6.62 ± 2.38 years) and 16 Children with Different Condition (CwDC) (10 boys and 6 girls; 7.75 ± 2.70 years) were recruited. Only 23 from ASD group and 15 from CG were included in the study, due to several reasons: four children did not finish the intervention due to anxious behaviors, two children declined to enter the music-therapy room and one did not meet the inclusion criteria.

4.3.2 Joint attention protocol

The robot-mediated intervention was designed using the Discrete Trial Teaching (DTT) protocol (SMITH, 2001), which consisted of the presentation of three types of stimuli: the antecedent stimulus (Sa), acceptable response (Ra), and consequent stimulus (Sc) to achieve a specific target behavior through a positive and progressive reinforcement system. In addition, the standard DTT protocol was complemented through a hierarchy of prompts. Thus, each triple of stimuli was composed of two different sequences of cues, which started from the least amount of help or least intrusive to the most amount of support or more intrusive prompts.

During the intervention, each participant came to the music-therapy room of HG clinic, and he/she was first exposed to a familiarization phase from three to five minutes. The therapist played with the robot and presented it to the child allowing to explore it. Parents accompanied the children, but they were instructed to avoid assisting the child during the intervention. Afterward, the first mediator therapist/robot began the JA task conducting the children's attention towards the objectives. Then, the mediator first greeted the participant ("Hi Pablo[3]. My name is Ono. I am a robot, do you like robots?... let's play") and provided the first level of antecedent stimulus, i.e., the mediator talks and looks toward the corresponding target ("Pablo, look at that toy!"). A successful JA event was defined as the child responding to the cue turning his/her head to look at the correct target within a five seconds interval. Then, if the child's response was the expected ($Ra =$ expected), the mediator rewarded the child with the corresponding Sc like ("Pablo, well done, let's continue!") and began the next trial towards the second target using the same stimulus level. Regardless of the child's response, the robot turned back to a neutral position after each prompt. If the child response was not the expected ($Ra \neq$ expected), the mediator repeated the sequence of Sa-Ra-Sc up to two more times using the first hierarchy level and then, changed to the next target. Once the three toys had been explored, the mediator proceeded to the following hierarchy level changing the stimuli from simple talking and looking to prompts combining pointing.

For the JA task, three toys (helicopter, motorcycle, truck) were located on the walls (left-front, right-front and right the child). The child together with the therapist were sitting facing each other, at a small table, situated in the middle of the room at an equal distance from the targets, while the robot was located on the table at left side of the child. Each intervention included three trials for toy at each prompt level (see Table 2), for a total of up to 36 trials across

3 No participant in this study has this name; this was chosen for illustrative purposes

Activity	Prompt's Hierarchy	Antecedent stimulus (Sa)*	Acceptable response (Ra)	Consequent stimulus (Sc)	
				Ra = expected	*Ra* ≠ expected
Introduction	Talking + looking toward the child	- *Hi friend, what's is your name* - *Hi Pablo. My name is Ono. I am a therapist/robot, do you like therapist/robots?* - *This room is amusing; we will observe together some objects that are in the room*	- The child responds with the name - The child responds verbally or makes a gesture of approval - The child responds verbally or makes a gesture of approval	- *Nice to meet*; wave with the arm - *Oh great!*; make happy facial expression - *Ok good!, let's play*; keep expression of happiness	- Repeat once and continue - Repeat once with neutral expression and continue - *let's play*
JA task 1	Looking + talking	- *Pablo, look at that helicopter* - *Now look at that motorcycle* - *Hey look at that cool truck*	- The child responds turning his/her head to look at the helicopter - The child responds turning his/her head to look at the motorcycle - The child responds turning his/her head to look at the truck	- *Pablo, well done, let's continue!* - *Good job!, one more time* - *Excellent!*	- *let's try again*; repeat up to two times and continue - *ok, one more time!*; repeat up to two times and continue - *let's again!*; Repeat up to two times and continue:
Transition	Talking + looking toward the child	- *Now let's see the toys again, but this time I'm going to point to those too*	- The child responds verbally or makes a gesture of approval	- *great!*; make happy expression	*Ok!*; make neutral expression
JA task 2	Looking + talking + pointing	- *Hey look at that nice helicopter* - *already saw that motorcycle, right?* - *Pablo, look at that great truck*	- The child turns his/her head to look or point toward the helicopter - The child turns his/her head to look or point toward the helicopter motorcycle - The child turn his/her head to look or point toward the helicopter truck	- *Well done, let's continue!* - *Good job!, one more time* - *Excellent, Pablo you were great*	- *let's try again*; repeat up to two times and continue - *Ok, one more time!*; repeat up to two times and continue - *let's again!*; Repeat up to two times and continue to the last activity
Closure	Talking + looking toward the child	- *It was a pleasure to meet you and play with you*	- The child responds verbally or makes a gesture of approval	- *Thank you and see you next time*; move the arms to shake hands	- *Thank you and see you next time*; move the arms to shake hands

* The Sa was administered randomly and the mediator held neutral expression

Table 2 – The structure of joint-attention task protocol.

Activity and Prompt's Hierarchy	JA performance score
JA task 1 - Looking + talking	2-Child turns his/her head or points after the first attempt of the robot/therapist 1-Child turns his/her head or points after the second or third attempt of the robot/therapist 0-Child does not react/does something else
JA task 2 - Looking + talking + pointing	2-Child turns his/her head and points after the first attempt of the robot/therapist 1-Child turns his/her head and points after the second or third attempt of the robot/therapist 0-Child does not react/does something else

Table 3 – Response scheme for assessing joint attention (JA) performance.

both therapist and robot intervention.

4.3.3 Children Behavioural Metrics

Choosing the correct metrics to describe the child's performance is essential to achieve a correct interpretation of the results. For the assessment scenario of JA impairments, two types of metrics have been identified: direct, and indirect (BEGUM *et al.*, 2016). The first one estimates the JA performance, computing a score of correct child's responses or quantity of JA behaviors. The second shows behavioral patterns associated with the task, such as visual contact, verbal utterances, goal oriented frequencies (LEMAIGNAN *et al.*, 2017), engagement (ESTEBAN *et al.*, 2017), energy and displacement (ANZALONE *et al.*, 2018), among others. The metrics used in this work are described below.

Joint attention score. The JA score was calculated using the behavioral grid presented in Table 3, which integrates child's behavior indicators, such as head orientation and pointing gestures as well as the time between the elicited cue and the child's response. Therefore, each child had three opportunities for each target to receive a score, and 3 targets along the room; in total, each child had 18 opportunities for mediator to exhibit a JA episode. Then, the JA score was rated according to the second column of Table 3. The maximum score is 12, which means that the child responded in the first trial by each target in the looking + talking task $(3 \times 2pts)$ and also he/she responded again in the first attempt to all cues for the looking + talking + pointing level $(3 \times 2pts)$

Adult seeking. Adult seeking behavior was characterized as looking towards the therapist or the parents immediately after an action performed by the robot mediator in the introduction script, where the robot greets the child and questions the child if he/she wants to play. This behavior was interpreted as a more sophisticated way of communication than the typical JA responses.

This behavior is composed of a combination of two kinds of JA response, i.e., the child exhibits an adult seeking event, when he/she tries to communicate emotion regarding the interaction (declarative behavior or protodeclarative) and at the same time the child manifests an approval seeking behavior to interact with the robot (imperative behavior or protoimperative) (STEINER *et al.*, 2012). The adult seeking event was considered valid only if this was exhibited into a three seconds window after the robot's cue. This metric was computed as the number of adult seeking events throughout the session.

Eye contact. An eye contact event was defined as three seconds interval of direct contact towards the therapist or robot mediator and was reported as the percentage that the child spent looking towards each of one.

Energy and displacement. The children behavior along the session was represented using the median energy associated with the rotation movement of the head and the median magnitude of head displacement. Also, these metrics were described graphically using a two-dimensional histogram of the movement and head orientation, where disperse histograms of ASD children are expected.

The energy of the head rotation $E(t)$ was calculated according to Equation 4.1:

$$\begin{aligned} E(t) &= \tfrac{1}{2} I \omega^2(t) \\ \text{with } I &= \tfrac{2}{5} M R^2, \end{aligned} \tag{4.1}$$

where ω is the angular velocity of the head; M is the head mass and R is the radius of the head circumference.

All metrics above were analyzed using the partial interval recording method. Thus, each intervention was analyzed using only an interval of 240 s period split as 120 s for therapist mediator and 120 s for robot mediator.

4.3.4 Experimental design and data analysis

A single exposure study used a 2×2 mixed-design with the interaction moderator (therapist TI or robot RI), as within-variable, and the condition CwASD and CwDC as between-variable. The dependent variable was the metrics described in Section 4.3.3. Half of the participants from each group was firstly exposed to the TI and then to the RI, and the other half conversely, in order to avoid potential order effects.

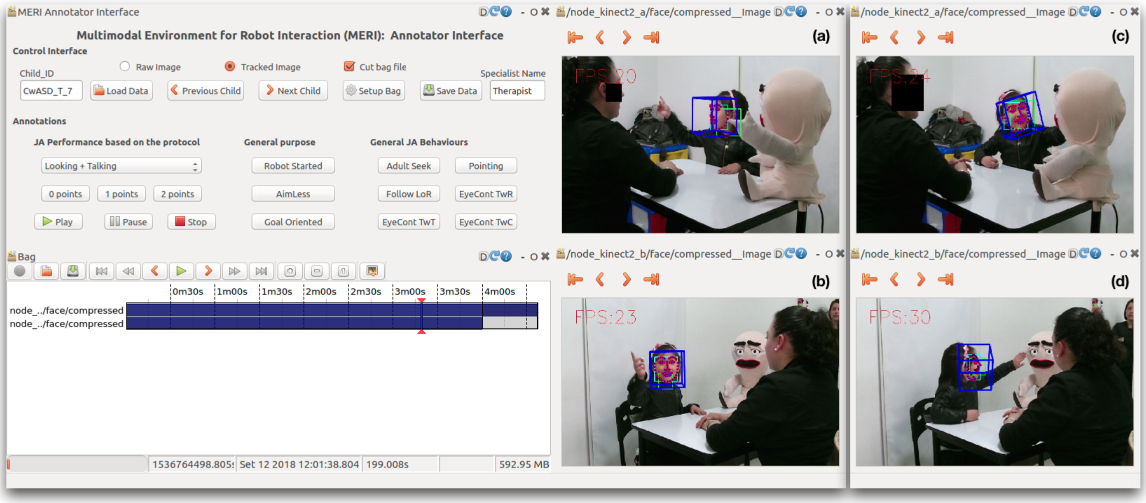

Figure 16 – MERI Interface appearance and video processing performance on a robot-based intervention. In the photographs, two behaviors that the children of both groups exhibited frequently can be observed: the child points toward targets (pictures a-b) and, touches the Ono's face (pictures c-d).

Statistical analyses were performed using SPSS version 25.0 (IBM, Armonk, NY, USA). Descriptive statistics were performed to describe the social and behavioral metrics. Due to the size of the sample population and the fact that the data is not normally distributed, non-parametric inferential statistical methods were used to analyze population differences and mediator effects. Thus, to test the potential as diagnosis assistive tool of the robot-based intervention, Mann–Whitney U-tests were performed to analyze the collected data regarding the different behavioral and social reactions using as between-subject factor the related children condition, i.e., CwASD vs CwDC. In addition, to compare the children's performance with a robot mediator and with the therapist, Wilcoxon signed-rank tests were performed using as within subjects factor the mediator, i.e., Robot-Mediated Intervention (RMI) vs Therapist-Mediated Intervention (TMI). An alpha level of 0.05 was employed for these analyses.

Two trainers who did not participate in the intervention design and protocol execution were trained in data coding. In the training process, the trainers selected and analyzed 20% of audiovisual recordings of the interventions with the collaboration of a researcher of this study. The final inter-rater reliability of the data coding revealed a Intraclass Correlation Coefficient (ICC) of 0.73 for JA score, 0.82 for eye contact and 0.85 for adult seeking metrics.

	Group	N	Mean	Std. Dev.	Mean Rank	∑ Ranks	M-W U	Sig. (2-tailed)
Head mdn energy	CwASD	23	0,00231	0,00217	18,04	415,00	139,000	0,329
	CwDC	15	0,00263	0,00185	21,73	326,00		
Mdn displacement	CwASD	23	0,40522	0,25708	20,17	464,00	157,000	0,658
	CwDC	15	0,34620	0,18950	18,47	277,00		
JA score	CwASD	23	9,6522	2,97885	17,98	413,50	137,500	0,275
	CwDC	15	11,3333	0,81650	21,83	327,50		

Table 4 – Comparisons of the outcome variables in the Therapist Mediated Intervention (TMI), according to groups.

4.4 Results of Robot-Assisted Intervention at Howard Gardner Clinic

Figure 16 shows the MERI Interface designed and implemented to execute all necessary actions in the robot-assisted diagnosis intervention. The video processing algorithm achieved a child's face tracking on 87% of frames in the 240 s session interval selected for analysis. The mean processing rate in frame per seconds (FPS) in all session was superior to 24 fps. In addition, the head orientation and VFOA were estimated in all video, thanks to the landmarks tracking stage, which compensated the missed frames on the face recognition stage. The automated metrics estimation of the MERI system was able to code the same three metrics coded for the trainers, achieving the following agreement percentage: 0.67 for JA score, 0.76 for eye contact and 0.79 for adult seeking.

The statistical analysis regarding the direct child's performance metrics showed that: with respect to JA score in the TMI scenario CwASD rating lowest than CwDC $(\text{ASD mean rank} = 17.98,\ \text{DC mean rank} = 21.83)$, however, the difference was not statistically significant $(U = 137.500; p = 0.275)$. The JA score in the RMI was significantly lower for CwASD than CwDC $(\text{ASD mean rank} = 15.63,\ \text{DC mean rank} = 25.43; U = 83.500; p = 0.005)$. This means that CwASD had greater difficulty in interpreting the robot's actions. In fact, CwASD performed significantly worse in RMI compared with in TMI $(Z = -3.629; p = 0.000)$. In contrast, CwDC rated better in RMI than in TMI $(Z = -1.823; p = 0.078)$. Some statistical results of

	Group	N	Mean	Std. Dev.	Mean Rank	∑ Ranks	M-W U	Sig. (2-tailed)
Head mdn energy	CwASD	23	0,00327	0,00186	19,61	451,00	170,000	0,953
	CwDC	15	0,00323	0,00202	19,33	290,00		
Mdn displacement	CwASD	23	0,64833	0,45822	22,00	506,00	115,000	0,089
	CwDC	15	0,45809	0,39927	15,67	235,00		
JA score	CwASD	23	5,1739	4,83032	15,63	359,50	83,500	0,005
	CwDC	15	9,6000	3,29068	25,43	381,50		

Table 5 – Comparisons of the outcome variables in the Robot Mediated Intervention (RMI), according to groups.

	Condition	N	Mean	Std. Dev.	Mean Rank	∑ Ranks	M-W U	Sig. (2-tailed)
Eye contact TwT (%)	CwASD	23	16,03	9,84	13,48	310,00	34,000	0,000
	CwDC	15	34,75	11,01	28,73	431,00		
Eye contact TwR (%)	CwASD	23	30,98	13,78	21,22	488,00	133,000	0,244
	CwDC	15	27,17	6,87	16,87	253,00		
Adult seeking events	CwASD	23	0,65	0,57	12,35	284,00	8.000	0,000
	CwDC	15	3,27	0,88	30,47	457,00		

Table 6 – Comparisons of the outcome variables throughout both intervention according to groups.

Mann–Whitney U-tests and Wilcoxon signed-rank tests are presented in Tables 4 - 6

The observations regarding the number of adults seeking events throughout the intervention, which is directly related to JA performance, revealed that in the RAD introduction CwDC manifested significantly more adult seeking events than CwASD $(\text{ASD mean rank} = 12.35,\ \text{DC mean rank} = 30.47; U = 8.000; p = 0.000)$. This fact can be interpreted as decreasing of interest to manifest pro-social behaviors, which are typically observed in CwASD. The details are presented in Table 6.

Regarding the visual contact towards the robot, CwASD exhibited slight more preferences to look toward the robot mediator compared with the results of CwDC $(\text{ASD mean rank} = 21.22,\ \text{DC mean rank} = 16.87; U = 133.000; p = 0.240)$. In contrast, CwASD showed significantly less visual contact toward the therapist than CwDC $(\text{ASD mean rank} = 13.48,\ \text{DC mean rank} = 28.73; U = 34.000; p = 0.000)$. Also, the difference of visual contact between mediators is statistically significant $(U = 56.500; p = 0.000)$. Thus, CwASD spent more time looking toward the robot than toward the therapist $(EcTwR > EcTwT)$ and CwDC exhibited the opposite pattern $(EcTwT > EcTwR)$. Statistical data are summarized in Table 6.

In relation to the head displacement and energy indicators, on the one hand the statistical analysis showed that CwDC $(\text{mean rank} = 21,73)$ spent more energy than CwASD $(\text{mean rank} = 18,04)$ in TMI, however, the difference was not statistically significant $(U = 139.000; p = 0.329)$. On the other hand, the analysis shows a interesting result, in the RMI both

	Condition	N	W S-R Z	Sig.(2-tailed)
Head mdn energy	CwASD	23	-2,859	0,003
	CwDC	15	-1,931	0,055
Mdn displacement	CwASD	23	-2,494	0,011
	CwDC	15	-1,022	0,330
JA score	CwASD	23	-3,629	0,000
	CwDC	15	-1,823	0,078

Table 7 – Comparisons of the outcome variables between therapist mediated vs robot mediated intervention.

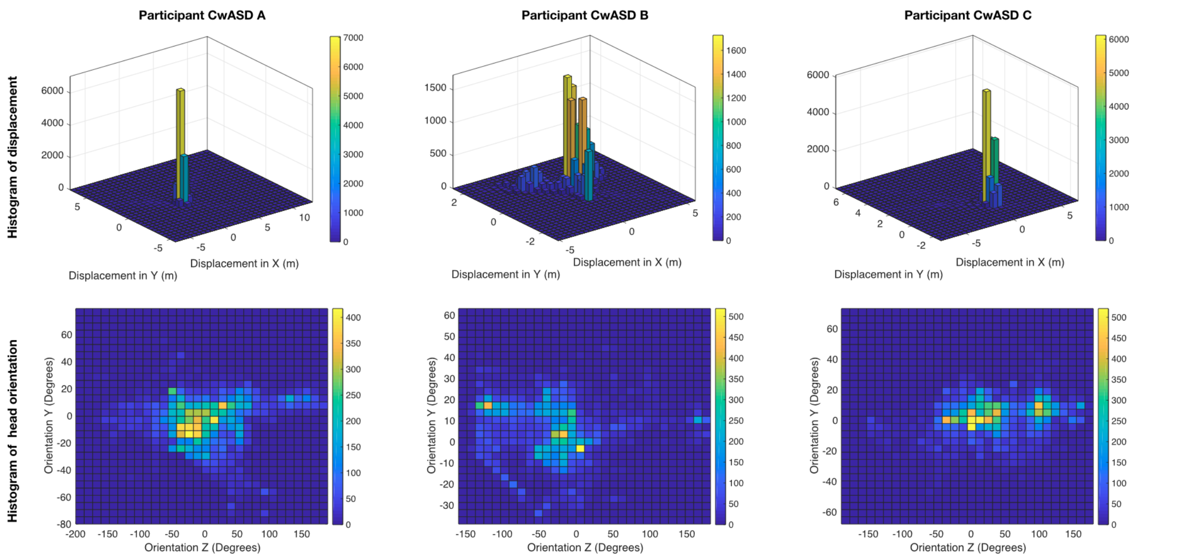

Figure 17 – Two-dimensional histogram of the movement and the head orientation for three CwASD.

groups exhibited a similar energy consumption pattern (ASD mean rank $= 19.61$, DC mean rank $= 19.33; U = 170.000, p = 0.953$). Thus, CwASD increased significantly their energy from the TMI to the RMI ($Z = -2.859; p = 0.003$), while CwDC reported similar energy in both interventions ($Z = -1.931; p = 0.055$). A similar behavior was evidenced regarding the displacement. The head displacement difference between groups was not statistically significant in TMI (ASD mean rank $= 20.17$, DC mean rank $= 18.47; U = 157.000; p = 0.658$) neither for RMI (ASD mean rank $= 22.00$, DC mean rank $= 15.67; U = 115.000; p = 0.089$). Also, CwASD increased significantly their displacement from the TMI to the RMI ($Z = -2.494; p = 0.011$) while CwDC did not change their head displacement ($Z = -1.022; p = 0.330$).

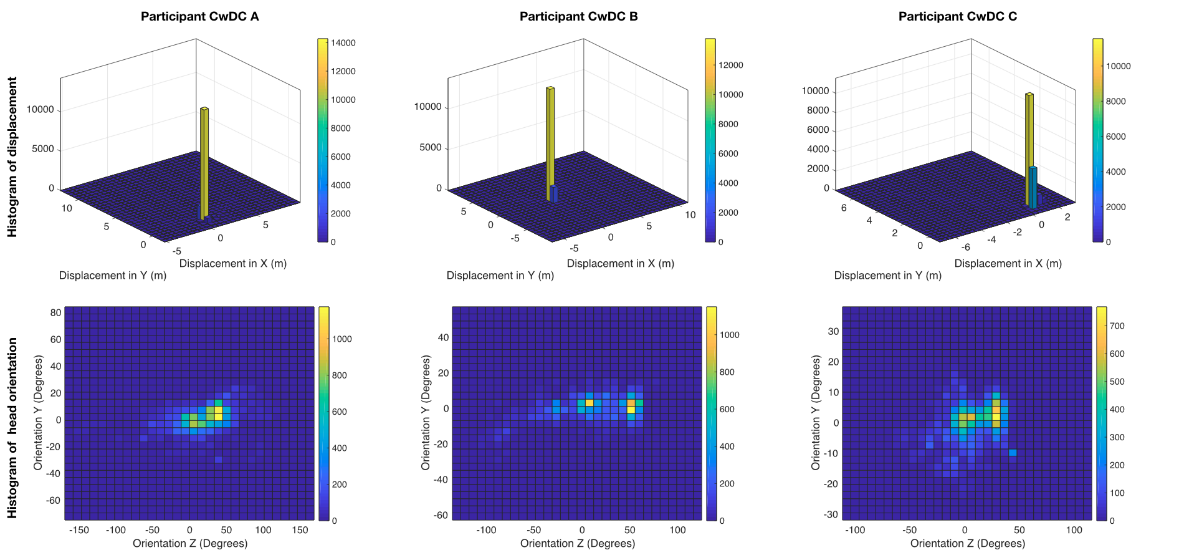

Figure 18 – Two-dimensional histogram of the movement and the head orientation for three CwDC.

In Figures 17 and 18, the total head displacement and rotation respect to the neutral position of three children for each group are shown. A significant difference in the behavior of the two studied groups can be also observed graphically. Thus, a comparison of both figures evidences that the displacement magnitude is higher for CwASD than for CwDC. The graphical analysis also showed a significant difference in the histogram of head orientation. In the case of the ASD group, the head's movement was higher, more dispersed, and less stable than in CwDC. These results can be interpreted as a measure of the postural stability of children, confirming that CwASD behavior in the JA task were less stable in terms of head movements in the environment.

Summarizing, CwASD exhibited more visual contact towards the robot mediator than towards the therapist, as well as, they show higher energy and displacement in the RMI. This could suggest that CwASD were more motivated to interact in the robot-mediated intervention. However, the increase in children's interest was not reflected in the children's performance, due to JA score in the RMI was the worst, and they did not manifest an increase of pro-social behaviors, such as adult seeking.


Steve Asbell

CHAPTER 5

FROM AUTOMATED DIAGNOSIS TO PERVASIVE DIAGNOSTIC PARADIGM

The use of information technologies and assistive robotic tools to improve traditional clinical processes employed for diagnosis and therapies of ASD has been increasing significantly in the last ten years (GRYNSZPAN *et al.*, 2014). However, such as aforementioned, technology-based applications for ASD therapies have emerged faster than applications for ASD diagnosis.

Technology-based intervention studies for CwASD include mobile-based applications, such as personal assistant (FRAUENBERGER *et al.*, 2011), virtual reality (SAADATZI *et al.*, 2018), serious games (BARTOLI *et al.*, 2014; BENTON *et al.*, 2014; MALINVERNI *et al.*, 2014), telemedicine (NAZNEEN *et al.*, 2015), wearable sensors for behavior monitoring (WASHINGTON *et al.*, 2017; NESS *et al.*, 2017) and human-robot interaction (BELPAEME *et al.*, 2012; SCASSELLATI *et al.*, 2012; VALLÈS-PERIS *et al.*, 2018; GOULART, 2019). As collective outcomes, studies have shown that technologies have the potential to aid ASD therapy and diagnostic process, though it is necessary to conduct more studies to support the efficacy of these tools. In addition, few applications tried to explore the best of each one to link some of these techniques in a common framework. This way, the aforementioned tools can enhance the diagnosis process and provide child and parents with early access to ASD diagnosis (ARESTI-BARTOLOME; GARCIA-ZAPIRAIN, 2014; GRYNSZPAN *et al.*, 2014).

This chapter provides a comprehensive theoretical scheme to show the potential role

This chapter was adjusted from: Ramírez-Duque, A., Bastos, T., Frizera-Neto, A., (2019). Multimodal Environment for Robot-Mediated Intervention (MERI): A New Paradigm for Pervasive ASD Diagnosis. *Manuscript submitted for publication.*

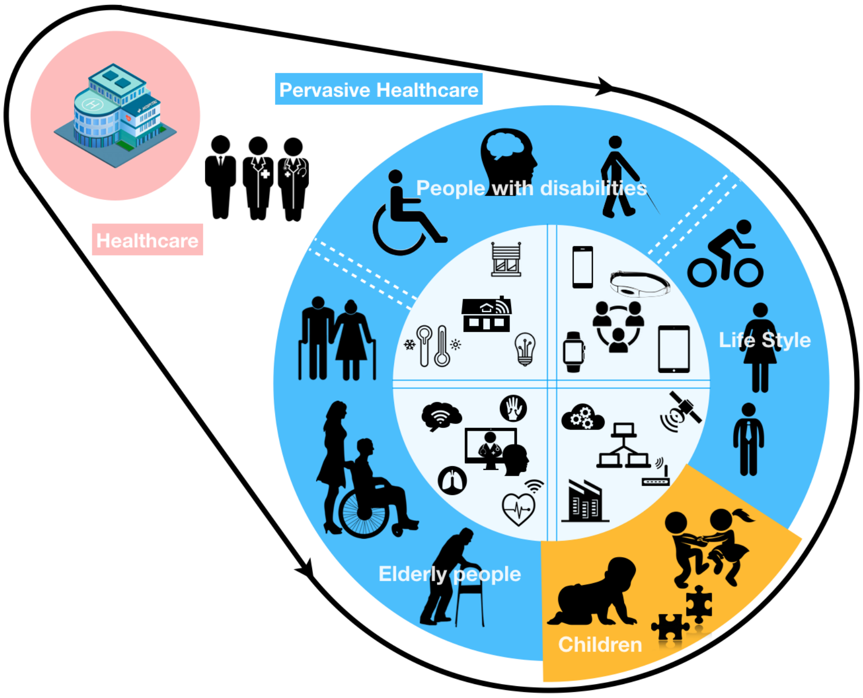

Figure 19 – Technological elements and population involved in a pervasive healthcare system. The target population of this research is highlighted in yellow.

that the integration of IoT technologies in the MERI system can play in the ASD diagnosis process. The theoretical approach leverages the combined strong synergy of IoT-based architecture, cloud computing and CRI to build a new paradigm for pervasive ASD diagnosis.

The growing healthcare perspective termed Pervasive Healthcare System (PHS) inspires our research. PHS is a strategy that seeks to provide a quality health system accessible to all people without constraints imposed by geographical location and time. PHS tends to gradually transfer traditional clinic-centered health systems into more personalized and mobile-centered healthcare systems, the well-known "healthcare to anyone, anytime, and anywhere" (VARSHNEY, 2007). The main technological elements and population involved in a PHS are illustrated in Figure 19. Further information about these techniques and tools can be found in (UNIYAL; RAYCHOUDHURY, 2014), and a survey of pervasive mobile healthcare systems can be found in (HUZOOREE *et al.*, 2017).

Currently, there are studies in which strategies of ASD diagnosis processes together technological approaches have been proposed, one of this is the telemedicine, whose procedure is a modern tool to assist the ASD diagnosis process remotely through the capture of home videos (KNUTSEN *et al.*, 2016; NESS *et al.*, 2017). In (NAZNEEN *et al.*, 2015) the authors presented two tools, NODA smartCapture, and NODA Connect, in which the former is a mobile phone-based application that enables parents to easily record clinically relevant prescribed video evidence of their child's behavior. The latter is a web-based application for clinicians to conduct a video tagging and a child's behavioral analysis. In addition, computational intelligence methods and machine learning models have been used to assist both ASD therapy and diagnosis process. In (THABTAH *et al.*, 2018) they presented an intelligence method called Variable Analysis to reduce features in three common screening methods. In (RUDOVIC *et al.*, 2018) the authors described a personalized machine learning for offline estimation of the children's emotional states and engagement during robot-assisted therapy. In (WASHINGTON *et al.*, 2017) they presented the "SuperpowerGlass" project to developing an at-home and parent-administered intervention for CwASD using the Google Glass.

Despite the potential that these technological-based techniques have shown, there are some clinical and technical challenges to be analyzed regarding their integration in a common framework. To the best of our knowledge, there are not many examples where telemedicine, cloud computing, and CRI are integrated, to apply pervasive healthcare strategies on ASD diagnosis process. Thus, the main contribution of this chapter can be summarized as follows: (i) MERI theoretical extension based on an unified layer design to allow cloud computing and data-driven; (ii) an alternative and feasible three-step process to identify the pathways from the first concern to pervasive ASD diagnosis.

5.1 MERI for Pervasive Autism Diagnosis: Challenges and Constraints

The primary challenge of this research is to develop a system through which it is possible to achieve the benefit of early ASD diagnosis that clinical community establishes. In practice, this means meeting rigorous clinical criteria in each of the stages considered to implement a new method or process, such as formulation, design, developing and validation through an in-clinic intervention. The clinical and technical challenges and constraints are further described in the following sections.

5.1.1 Clinical-Community Issues

Currently, both ASD therapy and diagnostic interventions have been developed in clinical environments, which in general, have been limited for factors such as available therapists, reduced time interventions and physical resources like assisted tools (WIGGINS *et al.*, 2015). These limitations can lead to inconsistent observations of the child's behavior. From the clinician point of view, observing the child's behavior in a natural environment is the most suitable method to diagnose autism, which is the greatest challenge of the proposed system. In fact, the comprehensive assessment done in the context of daily life can lead to an accurate ASD diagnosis (HUERTA; LORD, 2012).

Thus, a pervasive ASD diagnostic process has two practical challenges. The first one is to capture the child's behavior in natural conditions that allow identifying risk factors of ASD to establish evidence of the child's development with clinical validity. The second challenge is to provide a mechanism to allow the clinical staff to evaluate the evidence gathered systematically (NAZNEEN *et al.*, 2015). Besides, due the ASD diagnosis is comprised of multiple observation sessions, the clinical staff needs to know the child's behavior evolution among assessments.

As previously mentioned, many children are first recognized by their toddler care center. Accordingly, the pervasive ASD diagnostic process has an additional challenge to connect people and plausible places to perform a pervasive and ubiquitous functional behavior assessment such as schools, care centers and home.

5.1.2 Technological Issues

The main technological challenge revolves around the choice of the MERI architecture to ensure the system adaptability in different scenarios and achieve pervasive and ubiquitous performance. Thus, the MERI structure should focus on ensuring modularity, flexibility and scalability as well as connectivity and accessibility. In fact, the development of new elements to ensure a pervasive approach such as network design, data security and privacy, big data collection, processing and transfer is a challenge. Thus, in order to gather evidence at unconstrained environments, such as care centers and home, requires more complex sensing and computer routines are needed to deal with greater variation in physical conditions and child's behaviors.

The data storage and management represent two more challenges of the overall system. The data can help the specialists to observe behavioral patterns of the interacting child and track how these change over multiple intervention sessions and multiple environments. For the above, long-term remote monitoring system and appropriate feedback for clinical staff, parents and, developers must be supplied.

There is another challenge related to the level of autonomy and artificial intelligence that the system can exhibit. According to new trends associated with scalable machine learning (ML) evolution, some tasks can be performed automatically to optimize the overall system, which

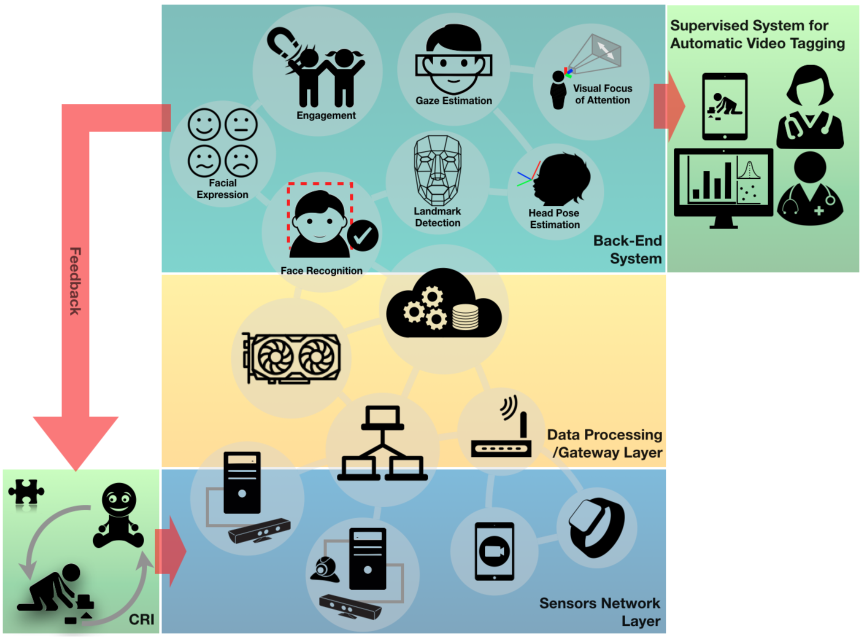

Figure 20 – IoT-based architecture of MERI system.

leads to more interoperable solutions and that can make effective decisions in the framework. Procedures that combines automatic perception of the children's emotion states, engagement (RUDOVIC *et al.*, 2018), and behaviors (SHORT *et al.*, 2017), producing the clinical history and generating reports (NESS *et al.*, 2017) are examples of the tasks that should be automated (ESTEBAN *et al.*, 2017).

5.2 The pervasive MERI

In order to deal with the main technological challenges, an IoT-based architecture was chosen to the proposed system. The connectivity and accessibility perspective of IoT architecture is well established because it allows creating an ecosystem that joins different elements via a cloud server (RAHMANI *et al.*, 2018). The cloud is responsible for data storage, computing, and visualization. In that sense, IoT architecture allows remote monitoring and tracking of children, in addition to create a continuum among plausible intervention scenarios through cloud access.

The MERI architecture extended here includes the following main components: sensor network, data processing, gateway layer and back-end system. A schematic diagram of the implemented structure is shown in Figure 20. The first layer is composed of a multi-camera array for behavioral analysis, and a wearable device for physiological signals monitoring. In the middle layer, there is a workstation network used for parallel processing and local data storage. The intermediate layer contains traditional gateways to connect mobile and wearable

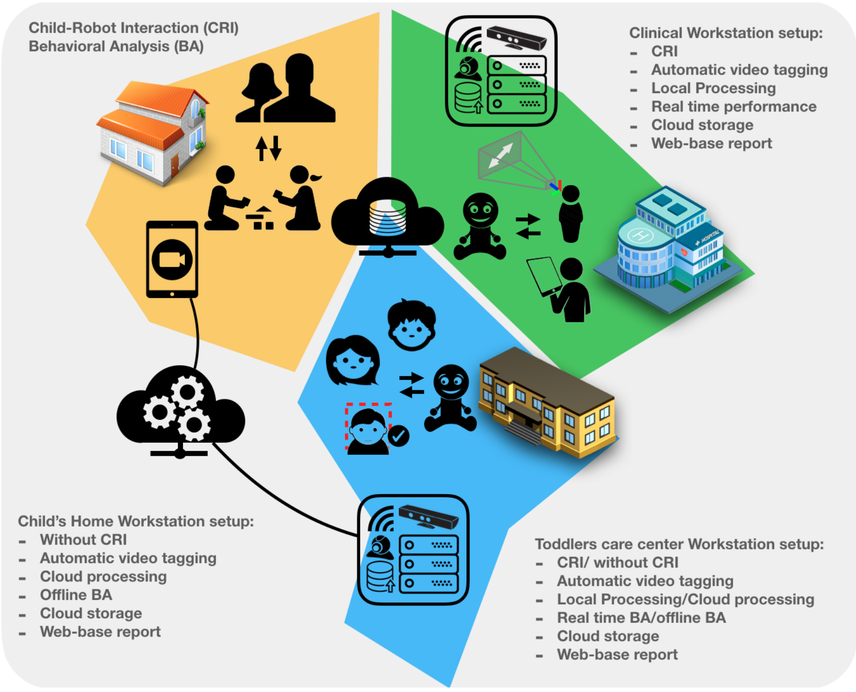

Figure 21 – A schematic of the novel autism diagnostic process using MERI system.

devices to the cloud container. In the upper layer, cloud computing and cloud storage models are implemented. The back-end solutions can be implemented using cloud platform based on OpenStack version Ocata and Amazon Web Services (AWS)[1]. OpenStack is an open-source cloud management software capable of providing and managing the network, processing, and storage resources in a data center. Virtual machines (VM) can be instantiated in the cloud to provide services for MERI. Furthermore, the architecture provides in the last level a responsive web-based applications set for end-users (parent, caregiver, and clinician) composed of graphical user interfaces for result visualization and clinical feedback.

5.3 Scenarios: MERI for Anyone, Anywhere and Anytime

In addition to pervasive MERI concept, in this chapter a feasible three-step process to identify the pathways from the first concern to ASD diagnosis is proposed in order to overcome accessibility and connectivity challenges. The proposed conceptual pathway is based on three plausible places where the community can be connected in a straightforward scheme. Thus, this section presents a novel strategy in order to develop a pervasive autism diagnosis process. All the steps are linked using the pervasive version of MERI.

In the conceptual pathway, three situations are defined to capture evidence of the child's behavior. The scenarios are places that children visit often and where atypical behaviors

[1] https://aws.amazon.com/

are most evident: the toddlers care center, the child's home, and clinical environment. The three process stages correspond to the capture of clinical evidence in each of the scenarios. A schematic of the novel autism diagnostic process using MERI is shown in Figure 21. The configuration of the MERI to adapt it to each one of the scenarios is described below.

Toddlers care center setup. The first stage of pervasive ASD diagnosis is realized as a collective intervention approach. The mediation consists of the implementation of a CRI protocol for group activities that allows evaluating the behavioral response of children to specific stimuli performed by the robot or caregivers. In the first instance, MERI is configured to collect intervention information about the children's performance using the multi-camera array. Data are analyzed independently for each child who participates in the session. Data parallel processing is performed through cloud parallel computing feature of MERI system. Pediatricians evaluate the automatic video tagging results and risk factor identification using the clinician-user session of the MERI web-based application. Then, a medical committee remotely assess the results of children's behavior. If some of the children reveal signs of atypical behavior, the clinician can refer the children to continue the next stage of the diagnosis process.

Child's home setup. Considering that children who reach this second stage may manifest signs that are common among different behavioral disorders, capturing more evidence to achieve conclusive results is necessary. The parents are oriented to record specific events regarding the child's day-to-day development. Also, the parents can complement the video information filling in a report about the child's behavior. In this scenario, the parents have access to the parent-user account in the MERI web-based application to do the upload of video captured and fill out the child's daily history. The app supports the most common video recording devices as mobile-phone and digital camera video recording. The web application also provides description and examples of the atypical child's behaviors considered as clinical evidence, as well as activities to stimulate them. The videos are processed in the cloud container and result in an automatic video tagging. After data processing, the videos and an automatically generated report are sent to the group of specialists to perform a supervised behavioral coding. The clinical staff accesses the outcomes through the clinician-user session web-application. In this case, the application allows adding or modifying the video tags and submitting an assessment report of the child's condition. According to the results obtained in this stage, specialists can suggest to the parents to continue or leave the process. The clinician feedback is leveraged in order to fit the learning process of automatic video tagging. It is necessary to highlight that in this autism diagnosis stage the MERI setup does not consider the CRI module.

Clinical setup. The last step is a functional behavior assessment using MERI in-clinic setup. In this intervention, results of the ASD risk factor identification are confirmed. In this configuration, again, MERI uses the CRI module to stimulate the behavioral patterns exhibition that enhances the clinical assessment. A clinical protocol that meets the requirements of evidence-based practice (EBP) in autism is also used. In the behavior analysis module, the fit model that result from the previous stages is also evaluated. Face detection, localization and recognition, child's intention, engagement and visual focus of attention estimation are executed. If the child tolerates the use of the wearable device, MERI may use the correlation of the physiological signals and the child's level of engagement. All data are processed in the local GPU boards of the workstation network for real-time response.

The last layer of MERI system is implemented with a security layer to ensure data protection and data privacy of participants, their families and clinicians through three services, cloud data encryption, user authentication and access authorization.

This novel strategy would close the cycle of pervasive autism diagnosis process used in different places, at different times and for any of the children who, once began intervention, have been identified with risk.

5.4 Beyond the theoretical concept

Analyzing the clinical setup of the robot-mediated intervention proposed in chapter 3 as a partial proof of concept of the pervasive MERI, some limitations were identified:

(i) In each intervention a high flow of data was generated to be transmitted and stored, given the resolution and the sampling rate of the captured videos. In a first validation, the captured videos were transferred and stored in the back-ended workstation. However, due to high data traffic the system lost processing capacity and exceeded the write speed of the hard drive. This fact led to the implementation of local storage and video compressing, which would be essential for the development of the future pilot test in other two scenarios that use both cloud storage and cloud computing.

(ii) The installation of the MERI in-clinic setup needed a large group of staff, which was a challenge. The above fact must be counteracted using compact network cabinet.

(iii) The open frameworks applied for the child's face analysis was trained with databases of adults' faces located near the camera. Conversely, MERI setup needed analyze the children's faces at a longer distance of the camera. Using a new database that meets these requirements, the performance of the module for automated child's behavior analysis can be improved.

From a general perspective, in each of the scenarios mentioned above, the pervasive MERI system may be used to extract different signs of ASD, such as eye contact, stereotyped movements of the head, concentration, excessive interest in objects or events, and the JA behaviors analyzed through this Thesis. Besides, given the advantages of layer-based architecture proposed and that data processing algorithms were based on state of the art deep learning models, more ASD signs can easily be analyzed by adding new deep models. Thus, the system could increase the number of ASD signs estimated as fast as new related neural models appear, and this is happening every day.

The concepts presented in this chapter emerged from a detailed analysis of different technology trends previously described as potentially useful to help the autism community. Therefore, the proposed system is based on a theoretical but ambitious concept that becomes more feasible every day.

In spite of the previously mentioned, demonstrating the practical validity of the proposed system is not an easy task. First, the consolidation of an extensive multidisciplinary program coordinated by governmental public health entities and with the support of academic and the autism community is required to develop any practical feasibility analysis. However, the evidence presented throughout this Thesis can be considered as a starting point to achieve this goal. Finally, in the near future, this proposal can transform the perspective of ASD diagnostic tools and thus contribute to bring children in ASD risk closer to the early diagnosis in a more efficient way.

Chapter 6

Collaborative and Inclusive Process with the Autism Community

The inclusion of social, cultural, and ethical concerns, in addition to clinical requirements to reach an evidence-based practice of RAT and RAD has been adopted in several research projects regarding CRI (RICHARDSON *et al.*, 2018; PECA *et al.*, 2016; COECKELBERGH *et al.*, 2016). In this sense, PD techniques have been explored to integrate the aforementioned aspects through the consideration of the contributions coming from all of the populations engaged in the process (e.g., stakeholders community) (FLETCHER-WATSON *et al.*, 2018). The use of PD methods in CRI research allows capturing the context and cultural factors in real-world situations, and shows tangible benefits when the research findings are moved into real clinical practices. The philosophy around a PD is to empower all of the people that are involved in a specific activity or situation (each of them in their particular level of involvement) by providing them space and voice, so that all they can contribute in the decision making during the process (GUHA *et al.*, 2014).

In this chapter, the development of a PD methodology is reported, which aims to identify guidelines for the design of a social robotic device to be implemented in a robot-mediated intervention for CwASD. PD is inherently reliant on the culture and views of the location in which it takes place and, in this context, the PD also represents an opportunity of gathering culture-specific findings and making cross-cultural observations in an ethical way.

This chapter was adjusted from: Ramírez-Duque, A., Aycardi, L., Villa, A., Munera, M., Bastos, T., Belpaeme, T., Frizera-Neto, A., Cifuentes, C., (2019). Collaborative and Inclusive Process with the Autism Community: A Case Study of Social Robot Design. *Manuscript submitted for publication.*

This study was developed as a part of the Compliant Soft Robotics (CASTOR) project, which aims to develop a compliant soft robot to build the next generation of ASD intervention scenarios.

From the existing procedures, a list of elements that could be incorporated into a PD method for the community around the ASD was selected. Thus, PD features that empower children and adults with ASD, as well as their parents, teachers, and caregivers were selected. Also, an increasing awareness has been placed on the importance of involving the community in the design process to better understand their wishes, concerns, needs, and preferences. There are a range of research projects, each using different methods. In these the main objective was to design technological tools as learning applications (FRAUENBERGER *et al.*, 2011), serious games (BENTON *et al.*, 2014; MALINVERNI *et al.*, 2014), interactive environments (MERTER; HASIRCI, 2016) and robotic-devices (BERTEL *et al.*, 2013; ZUBRYCKI *et al.*, 2016; HUIJNEN *et al.*, 2016b; VALLÈS-PERIS *et al.*, 2018). Regarding the participatory methods used, most of the projects implemented activities that allowed the children and stakeholders to participate in different roles, engaging as users, testers, informants, and designer, thereby increasing their motivation and their ownership of the project.

In general, the two first roles –users and testers– were the most commonly used, even though they are considered to allow only *passive participation*. The information most often was obtained through questionnaires and interviews with parents and teachers, and sometimes through using observations of the children's behavior before and after using the designed object (MERTER; HASIRCI, 2016). In other cases, parents and clinicians were involved both as informants and designers, allowing more participatory and deciding roles. For example (i) for the design the intervention protocol in a particular school for children with ASD using the robot seal Paro (BERTEL *et al.*, 2013); (ii) the design of a robot-based environment to support the therapy for severe CwASD (ZUBRYCKI *et al.*, 2016); (iii) the definition of the role and some aesthetic features of the Kaspar robot to interact with CwASD (HUIJNEN *et al.*, 2016a; HUIJNEN *et al.*, 2017). On the other hand, a remarkable effort was in (HUIJNEN *et al.*, 2016b) where authors consolidated and proposed a list of the main domains and objectives where SAR could be implemented in a CRI to strengthen ASD therapies. In this case, the approach was performed through the involvement of focus groups.

The active participation of children, while often straightforward, can become challenging when the children have special needs. However, different participatory methods which

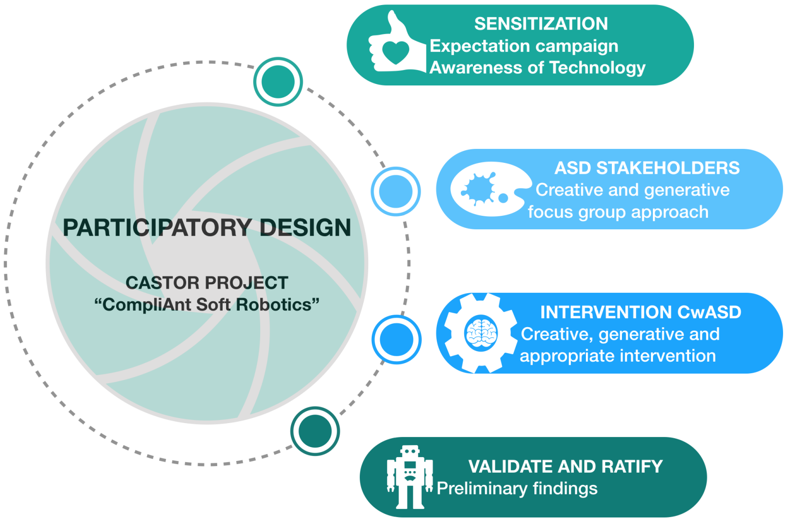

Figure 22 – General scheme of the Participatory Design (PD) in the first year.

involve children have been reported (HENDRIKS *et al.*, 2015). A common strategy was based on providing narrative structures to develop a story, stimulating curiosity and inviting and encouraging children to contribute. For example, in (FRAUENBERGER *et al.*, 2011) the authors provided high-functioning CwASD with a story based on a comic strip. The authors gave them the start and end points of the comic and asked the children to incorporate a particular object available during the activity. In (MALINVERNI *et al.*, 2014) they used scene cards to guide a narrative task and relied on drawing activities to develop video-game characters. In (BENTON *et al.*, 2014) the authors used a generation process based on a visual template, using physically drawing and art materials to allow CwASD to design a math-based computer game. Finally, in (VALLÈS-PERIS *et al.*, 2018) the authors used narrative-based participatory methods using free drawing activities, modeling paste and construction blocks sessions to analyze the children's view of HRI in health care.

In this work, the goal was to explore the benefits brought by the use of PD methods in the design of a social robot, with a specific focus on their use in ASD interventions. Based on what proved to be effective in earlier work, participatory methods for both the CwASD and the stakeholders were implemented. The process used a focus group approach with parents and therapist, and used scene cards, narrative and handmade generative methods with the children.

6.1 Methods

Implementing PD is not just a methodology to improve and enhance a product's final design, but also an opportunity to understand and gain knowledge about the community's context, and to build trust and confidence between researchers and the community. It is also a chance, in this case, to show the benefits of technological tools in this complex social context. Thus, the participatory process was made up of different stages that could lead us to achieve the following objectives:

- obtain contextual information that allows the establishment of the needs, interests, preferences, fears, desires and priorities related to functionality of the robot and its use as a tool in ASD therapies;
- validate the insights gained through the literature review for the design of the robotic device;
- generate ideas and creative solutions through reflecting on our experiences;
- promote the take-up of our research process and its results.

The process for designing a compliant social robot appropriate to be used in ASD therapies with children was planned together an interdisciplinary team. The CASTOR team included the creative enterprise specialized in inclusiveness design “Tejido de Sueños” [1], a Howard Gardner Clinic group composed of healthcare and administrative specialists and an engineering group. According to the agreements reached in the work sessions of the CASTOR team, the participatory process was established as a two year-long project. The first year was planned into four phases: (i) sensitization; (ii) focus group with stakeholders; (iii) generative intervention with children; (iv) validation and ratification of the preliminary findings. The implementation scheme for the first year is illustrated in Figure 22. For the second year, four more stages were planned: (v) perceptual maps and conceptual design; (vi) preliminary 2D/3D prototyping with community feedback; (vii) detailed design and manufacturing, (viii) presentation and validation of results. A summary of all CASTOR’s phases with the objectives in each stage and proposed activities is schematized in Table 8. The focus of this work is to report the findings of the first year of the CASTOR project. Therefore, a detailed description of the methodology and main results of the first four phases is described below.

[1] Design company for inclusion: https://www.tejidodesuenos.com/la-empresa

Year	Phases in the participatory design	Objectives to be reached	Activities to perform with:	
			Children	Parents and therapist
1	(i). Sensitization	• Introduce the context, the objectives and the team of the CASTOR project in the ASD community as well as identify the potential areas for robot intervention and motivate the community to participate in the participatory process	Workshop session Robot intervention demonstration	Workshop session Robot intervention demonstration Information support regarding CASTOR project
1	(ii). Focus group with stakeholders	• Inquire about the needs and opportunities to improve ASD therapies • Identify interests, wishes, and fears regarding the use of a robot to assist the specialists in therapy processes • Generate ideas and proposals for the robot design and the potential benefits to ASD therapies through the use of a robotic-device	Creative robot design using recycled materials, free play and dancing	Therapy context mapping CRI Idealization Creative robot design Domain intervention prioritization
1	(iii). Generative intervention with children	• Identify how the robot's physical features and interaction interfaces influence the perception of CwASD • Identify the children's preferences for different morphological characteristics of the robot • Explore the thoughts and preferences of children about the design of a robot based on didactic exercises	Sorting cards Adjectives matching Morphological variants Design my robot friend	NA
1	(iv). Validation and ratification of the preliminary findings	• Validate and ratify preliminary design decisions and plausible intervention domains for SAR • Collect additional information about the acceptability of the developed activities	NA	Personal interview Online questionnaire
2	(v). Perceptual maps and conceptual design	• Establish design requirements and conduct requirements analysis based on collected evidence through primary and secondary source	Activity to be developed by the CASTOR members	
2	(vi). Preliminary 2D/3D prototyping with community feedback	• Propose many sketches, diagrams and prototypes of the proposed robot to be tested with different actors	Storytelling based on proposed sketches Drawing activities: child's imaginary and prototypes idealization	Workshop: iterative validation of prototypes' feature and final decisions regarding the robot design
2	(vii). Detailed design and manufacturing	• Choose materials, sensors, passive and active elements, detail dimensions, assemblies, interfaces • Build the robotic-device based on the community's preferences and needs	Robot-mediated intervention demonstration in laboratory setup using the final prototype	NA
2	(viii). Results: presentation and validation	• Present the project's outcome and validate the robot through the use in robot-assisted therapy based on the community's demand for CwASD in the HG clinic	Implementation of a well-established protocol for robot intervention in clinic-setup	Workshop: CASTOR's outcome analysis and feedback throughout two-long years process

Table 8 – Summary description of the two year-long participatory process for the design a social robot with the participation of autism community.

6.1.1 Sensitization

The sensitization phase aimed at introducing the context, the objectives and the team of the CASTOR project to the local ASD community. Likewise, the sensitization phase allowed the CASTOR team to learn about ASD and interventions, the personal experiences of parents and specialists, in addition to the needs and concerns of the ASD community. This phase comprised two steps. The first one consisted of an expectations campaign, in which we talked about the project over several visits to the clinic. Throughout this time several activities were carried out to query parents and stakeholders about their views and ideas on and about robotics. Initially, a drawing of a robot was displayed in the facilities of the clinic along with a mailbox with three questions. The first question asked stakeholders to describe how they imagine a robotic device. The second question enquired about how robots could assist in therapies. The third question asked the stakeholders about how they imagined a robot to benefit CwASD.

Over those two weeks we organized several robot demonstrations and showed videos about different types of robots. The robot demonstrations aims to exhibit in a practical context the main elements that compose a CRI scenario. Thus, the music-therapy room in the HG clinic was equipped with two robotic-devices: Nao and ONO. The music-therapist addressed all session and only the essential functions of each robot were explained. The two robots were controlled by a CASTOR engineer using the *Wizard of Oz* approach. The engineer followed the interactive actions that the therapist determined throughout the session spontaneously, suiting each child's preferences. The robot-based demonstration had the only particular objective of presenting the CRI potential to analyze the feedback of the therapists, with no therapeutic objective considered. Finally, the expectation campaign was closed by a formal presentation about the two-year CASTOR project.

6.1.2 Focus Group with Stakeholders

The focus group phase first served to build a common ground between researchers and the community, and allowed sharing experiences and views about the role of each actor involved in the process. The first activity relied on a form which was sent out before the session to parents and therapist in order to do a customized context mapping. In the form, some personal aspects and expectations about the activity were asked, and four questions were included to inquire about the needs and opportunities to improve ASD therapies as well as to identify current

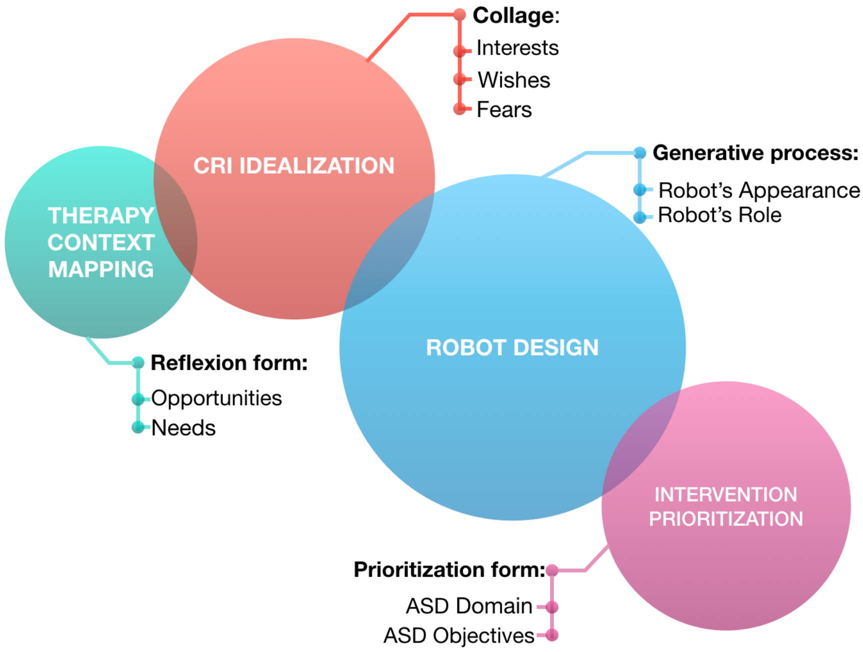

Figure 23 – Diagram of focus group developed with parents and therapist.

and future interests, wishes and concerns regarding the use of a robot to assist specialists during therapy. The four questions are described as follow:

- What are the positive and negative aspects of current ASD therapies?
- If you could create a magic tool to help therapy, what would this tool be?
- Imagine a therapy with robots, what would be the best and the worst aspects?
- Imagine that you have used the therapies with robots for more than ten years. In this case, which is the best or the worst thing that has happened to you so far?

The participants had small-group discussions about their opinions on all four questions. Then, in the second step, a member of the CASTOR team who acted as moderator asked the participants to imagine the ideal robot intervention and describe this through a collage. The moderator invited the participants to include emotional aspects, actions and impressions. The session ended with a plenary discussion, giving each participant the opportunity to express their views to the others.

The main generative step was implemented as an unconstrained creativity activity using recycled materials. The participants were placed in small groups and requested to create a robot. A final plenary discussion served to formulate guidelines for researchers and designers.

Finally, the prioritization domain activity closed the focus group session. This last activity aimed to elicit the five most essential domains or objectives for robot-assisted therapy for CwASD. The objectives and domains used correspond to those identified in (HUIJNEN *et al.*, 2016a). A summary of the activities implemented in this phase is illustrated in Figure 23.

6.1.3 Generative Task Developed for CwASD

The aim of this phase was to provide CwASD the opportunity to actively participate in the decisions that affect them. With that idea, four simple generative activities were designed considering the condition of the participants (see Figure 24). A set of six cards of the same size were prepared with references/images of robots commonly used in ASD therapies. In order to avoid aesthetic bias, the robot cards were selected as follow: (a) two anthropomorphic robots; Kaspar (ROBINS; DAUTENHAHN, 2014) and Nao (BELPAEME *et al.*, 2012); (b) two biomimetic robots; Probo (SALDIEN *et al.*, 2010) and Pleo (KIM *et al.*, 2012) and (c) two non-biomimetic robots; Leka[2] and Romibo (SHICK, 2013). The set of cards is illustrated in Figure 25.

In the first stage, all six cards are placed on a table in front of the child. After telling a short story about all the robots, the therapist asks the child to take the card that she likes the most. Once the child chooses a robot, the corresponding card is removed while the other cards are kept on the table. The therapist continues asking the child robot she likes most, and the sequences is repeated until all robots are ranked. In the second stage, the cards are again placed in front of the child, and another set of cards depicting different adjectives are laid out as well. The second set of cards shows six adjectives by using the following emoticons/pictographs: *cute,*

[2] https://leka.io/en/index.html

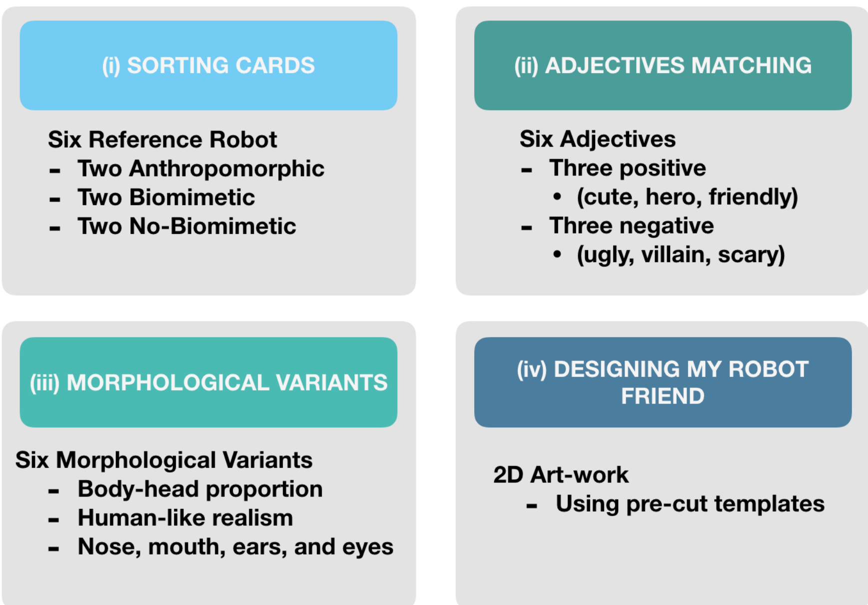

Figure 24 – Scheme of the generative task developed for CwASD.

ugly, hero, villain, friendly and scary. The therapist invites the child to match each robot with an adjective, this allows the therapist to see which associations or feelings the child has for each robot. During the third stage, the therapist presents the child with a third set of cards with aesthetic modifications of three robots. This serves to learn the child's preference for specific robot features. For each robot category (anthropomorphic, biomimetic, non-biomimetic) we chose a robot where particular aesthetic or physical traits were more prominent. In this way, Probo's image was chosen to modify the size of the ears and the body-head proportion. Romibo's image was chosen to show the modification of mouth and eye size. Kaspar was chosen to modify the human-likeness and the nose size (see Figure 25). In the fourth and final stage, the therapist asks the child to make a collage of their robot by choosing different pre-cut templates of heads, bodies, eyes, mouths, and noses, representing different morphological characteristics. At the same time the therapist encourages the child to comment on their decisions while building the robot. This activity is also used as an incentive for their participation, as the children can keep the art work if they so wish.

6.1.4 Validation and Ratification of the Preliminary Findings

In this phase, a questionnaire for the ASD community was designed to confirm and validate the findings of the previous phases. The questionnaire was distributed in the Howard

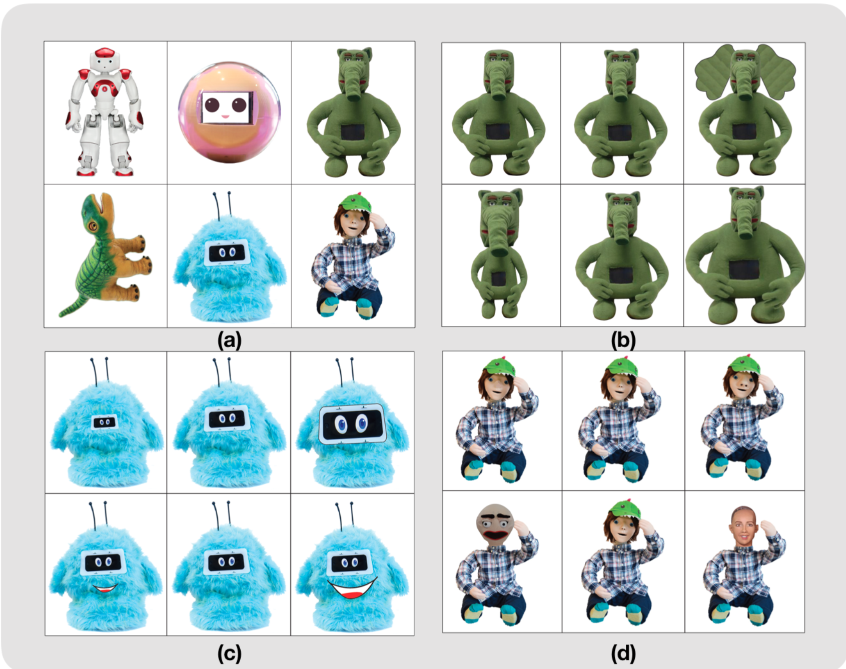

Figure 25 – (a) Set of cards with six robots commonly used in ASD therapies; (b) Aesthetic modification of the robots Probo; (c) Romibo and (d) Kaspar.

Gardner clinic and through social networks to other Colombian institutions specialized in the treatment and development of ASD therapies. It consisted of nine items related to the robot's physical features, 17 items related to the robot's physical behavior, one item about the use of sensory elements, six open questions about the role of the robot in the intervention and nine more general questions.

6.2 Results

The data generated during each phase of the participatory process were analyzed. For this CASTOR team observed video recordings allowing them to understand the activity's atmosphere and identify the main aspects. The stories, attitudes, and opinions that appear repeatedly, the surprise comments, the novel concepts that were uncovered, and the positive or negative responses of the participants were noted. Four members of the research team independently transcribed the recording, taking into account the primary purpose of the activity. The individual results were discussed, and a final summary was generated to gather the most important findings.

6.2.1 Sensitization

The sensitization phase took place in the two HG headquarters, one week was spent in the first headquarter, followed by a second week in the second headquarter. The robot art work and the mailbox were placed near the clinic's lobby, so children and stakeholders would engage. Additionally, during the sensitization process, a member of the CASTOR team invited the therapists, parents and caregivers that were in the clinic to participate in the official workshop introduction and the focus group activities. CASTOR members spent a considerable amount of time promoting the workshop.

In this phase, 18 stakeholders responded to three questions posted by the mailbox. Regarding the first question 61% of the respondents associated words such as *apparatus*, *machine* and *tool* to describe the robot. The other 39% of the participants assigned to the robot abilities to perform automated functions, in addition to artificial reasoning. In the second question, 17 participants answered that they believed that robot could be suitable and very useful to help CwASD, and only one person answered that he/she did not know about the subject. Finally, regarding the third question the most used phrases to describe the possible benefits of using robots in therapy were to *help them communicate*, *keep them motivated in therapy* and *teach them to play*.

Three CwASD, eight parents and four therapists actively participated in the workshop. During this , the participants expressed their interest to participate in the whole process and highlighted the importance of carrying out this type of activities together with the community.

Regarding the robot demonstration, a total of six CwASD and their parents participated in the activity. Furthermore, 12 therapists were spectators of at least one session. Some exciting results about the children's reaction and how the music-therapist conducted the session were perceived. First, the six children demonstrated high interest in both robots, without showing any preference patterns. Additionally, the music-therapist noted that all children showed increased pro-social behaviors, such as joint attention, imitation and verbal utterances, compared to the standard music-therapy session. Despite that music-therapist was not trained with regards to the robot's interaction capabilities, she was able to actively include the robotic-device and improve the child's attention and motivation in the six sessions. Interventions were characterized by the therapist's experiences about each children's condition and the therapist's creativity to engage the child in the robot-based play. The therapist started the interaction exploring proprioceptive and body awareness tasks. In addition, the therapist leveraged the salient features of each robot to improve the interaction. For example, the therapist used the Ono's facial features to perform facial gesture imitation, and the Nao's morphology to do free play and dancing.

6.2.2 Focus group

In the HG Clinic four focus group were set up, two for parents and two for specialists (therapists and caregivers). In each group, the same four stages were used, but adapted to consider the relationship the participants had to CwASD. The focus groups were organized in the clinic's facilities taking into account the stakeholders' availability. A total of 14 parents (N=14, all female, no age data available) and 16 specialists (N = 16, two male and 14 female, average age, 24 years) of the HG Clinic participated in the design process. The specialists had worked for at least two years with children with a variety of impairments, including ASD, intellectual disabilities, learning problems, and cerebral palsy. Both parents and specialists reported no previous experiences with any robot or related robotic-based activities.

According to what stakeholders expressed in the focus group, three main issues of including robots in ASD therapies are summarized as main negative aspects: adverse emotional

Figure 26 – Robots designed in the generative session of focus group.

reactions of the child, negative conditioning of the child’s behavior and loss of human factors in the therapies. However, many positive aspects were also identified. For example, both parents and specialists agreed that through robot-assisted therapy they would wish to increase the child’s motivation, reduce the child’s anxiety, improve the understanding of the child’s emotion, and enhance the child’s confidence in therapy. The findings of these aspects will be extended in the discussion section.

In the generative activity, all participants described a robotic device that could be composed of colored lights, different textures and materials, have buttons and a screen, have clothes as well as a face, upper limbs, microphone and speakers to communicate and interact with the child. Some robot prototypes specifically built for the participants are shown in Figure 26. Regarding the size, the participants expressed that it would be desirable that the robot’s were located at the same height as the child’s to facilitate the interaction. In addition, they think that materials and structure used in the robotic devices should allow physical contact, such as hugging and shaking hands

6.2.3 Generative Task with CwASD

The activity designed for children was run by a psychologist at HG clinic. A total of 11 CwASD (three female and eight male) with ages between three and nine years (5.81 ± 2.08) participated in the activity during their psychology sessions. The event took a maximum of 20 minutes.

Throughout the activity, the CwASD exhibited varied preferences regarding the

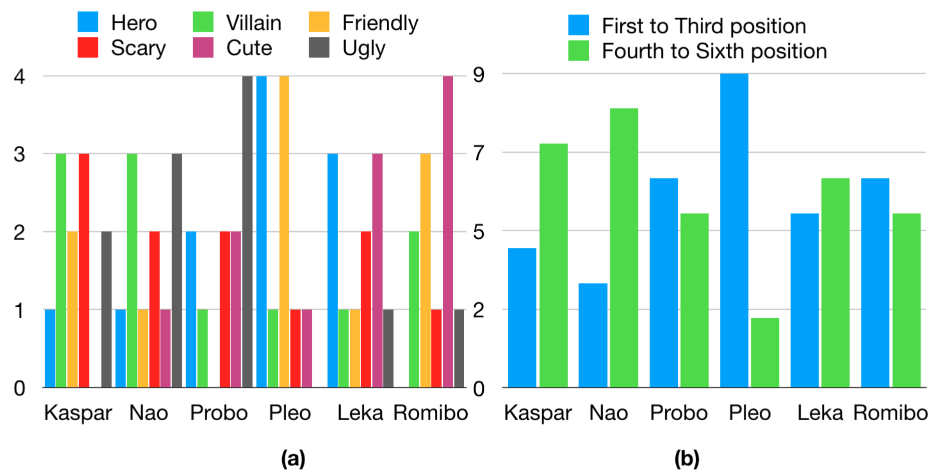

Figure 27 – Graphical representation of the results in the (a) adjectives matching; and (b) sorting card activities.

robot's appearance; however, some agreement emerged. For example, in the card sorting activity, nine children chose the Pleo card within the first three positions, followed by Probo and Romibo cards, which were selected six times within the first three positions. In contrast, the Nao card was chosen eight times within the last three positions, followed by Kaspar, which was chosen seven times.

A similar pattern was evident in the adjective matching activities, in which 82% of children assigned positive adjectives to Pleo, and 73% matched negative adjectives to Kaspar and Nao. In summary, four children assigned the hero adjective to Pleo while four other children described it as friendly. Romibo, with four votes, was chosen as the cutest, followed by Leka, with three votes. Four children assigned the ugly label for Probo, making him the ugliest. Regarding the villain adjective, Nao and Kaspar obtained the maximum score, with three votes each. Kaspar was also chosen as the scariest with three votes. A summary of the results is shown in Figure 27.

A general analysis of the last activity showed that the 11 children had preferences for exaggerated facial traits, such as a large mouth, ears, and nose as well as large and expressive eyes. Finally, in the generative activity, five children chose a dragon body, and animal-like heads for the robot; four chose a robotic body with a biomimetic head (two animal heads and two human heads), and two picked a human body with a robotic head. All children showed motivation during the creative activity; some of them showed an increase in communication using words and non-verbal signs to express enthusiasm regarding their final sketch.

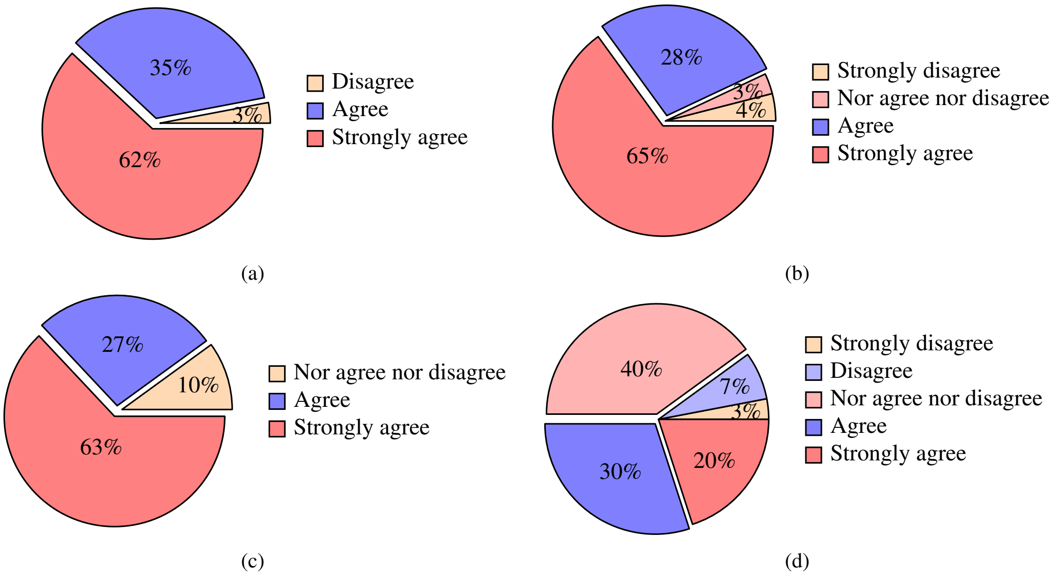

Figure 28 – Design decisions with regards to the robot's physical features and behaviors. (a) "it is convenient to add items such as clothes and accessories to customize the robot". (b) "the robot must be composed of a head, trunk, arms, and legs". (c) "the robot's facial expression must be similar to the humans' facial expressions". (d) "the robots must appear like a fantastic animal or character".

6.2.4 Validation and Ratification of the Preliminary Findings

A questionnaire used to confirm (or reject) our preliminary findings was distributed to parents and therapists in the HG Clinic. Direct interviews with members of CASTOR team were also used. In total 30 stakeholders (14 relatives of CwASD and 16 specialists) participated in this last stage.

The results showed a preference for the use of modular parts, plastic and textile materials, and a soft body. The participants believed that acoustic functions, movement of arms and facial expressions (mouth, eyes, and eyebrow movements) are essential for the CRI scenarios. Additionally, 97% of the surveyed agreed with the following expression; "it should be possible to add items, such as clothes and accessories to customize the robot," and 93 % of the participants agreed with the statement that "the robot must be composed of a head, trunk, arms, and legs". Regarding the statement "the robot's facial expression must be similar to human facial expressions", 90% of surveyed agreed, and with the sentence "the robot's eyes should be at the same height as the child's eyes", 87% of the respondents supported this. The participants were asked about the robot's appearance, 50% agreed that "the robots must look like a fantastic character", while 40% expressed neither agreement nor disagreement. The summary of the aforementioned results is shown in the Figure 28.

Figure 29 – Activities developed in the four focus group (two for parents and two for therapists).

6.3 Community awareness after PD

Through the implementation of the first four phases of PD process, a relationship of trust and understanding was established between parents, therapists and researchers. This was essential in this first steps of the CASTOR project, in order to balance the power distribution between the different actors and to assure a productive process. It is important to highlight the relevance that this aspect has for later stages of the process. At the beginning of CASTOR project, when we were looking for a collaborative partnerships, the first responses by health care staff included sentences like *"We are not willing to participate because researchers always use us to collect data and never come back; this is something that therapists and parents do not like"*.

Thus, through our collaborative and inclusive approach, the project has emphasized the necessity of prioritizing the people's well-being and community awareness rather than the technology results of this type of participatory process. The activities implemented also became an opportunity to spend time with other partners in a new context that could hold all people involved. In Figure 29, four pictures of the developed activities with parents and therapist are presented.

6.3.1 Awareness of CASTOR project in the sensitization phase

In the sensitization phase, the parents and therapists had the opportunity to know and understand the purpose of the CASTOR project, learning about the evidence gathered in other countries and regions about the potential benefit of the robot-assisted intervention. Through our expectation management campaign, a positive environment was created and, for this reason, all participants were more open to participate in the creative and generative activities. Furthermore, the participation and positive attitude towards the project contributed to the building of an enjoyable and collaborative atmosphere between volunteers and CASTOR team. As participants progressed between the activities and phases, both parents and therapist felt uninhibited to express their opinions and creative ideas, positively enriching the outcome of the process.

Despite the weak influence of robotic-based technology in a low-income country such as Colombia, the findings of the expectation campaign showed that the participants appear to be in favor of using robots within the therapies.

6.3.2 Findings from and reflections on focus group discussions

Throughout the four groups, many interesting opinions and ideas were generated, but only the commonly agreed upon ideas are described below.

With respect to the current needs of autism spectrum therapy, stakeholders expressed that it is necessary to modulate child's behavior before a therapy session. The route from their houses to the clinic or an external event before coming to the clinic can alter child's behavior and thus waste the therapy time due to, for example, anxiety. In the above scenario, the robotic device could help with modulating the child's behavior and reduce anxiety through free play or music. Thus, participants suggested that the robot should be equipped with sounds of familiar animals, musical instruments and songs. Parents expressed the same concern about child's behaviour and agreed that the robot could contribute to reducing anxiety, even in home, through using peaceful sounds.

"The robot could reduce child's anxiety levels, transmit calmness as well as confidence and thus avoid crises."

In addition, in the plenary discussion stakeholders acknowledged that child's motivation can be increased through robot-assisted therapies, and that this can be used to help the child in other aspects, such as communication and activities of daily life. Also, participants highlighted that with the robotic platform it is possible to stimulate the child using different communication channels, such as visual, auditory and tactile senses, proprioception and spatial exploration.

While many positive aspects emerged in the focus group discussions, participants also expressed some concerns. For example, both parents and therapists agreed that a major concern is that in the medium to long-term, child's behavior could be conditioned for the robot presence in therapy, i.e., the child can find so much comfort in interacting with the robot that later on, he or she will not want to interact with anyone else. Related to this one of the parents expressed:

"The robot therapy could limit my child's imagination and behavior; it could condition to the point of imitating and preferring the voice of the robot, getting used to the robot until he will not want to interact with other people."

Two more concerns emerged during focus group discussions. The first one was related to whether robot-assisted therapy can generate stress, anxiety, and frustration for the child, due to abrupt movements of robot, very loud or strange noises, or sudden mechanical failures. The second one, even though the focus group moderator reassured that a therapist would always facilitate therapy, both parents and therapists expressed concerns about reduction of human contact and the reduced amount of human emotional contact when using robot-assisted therapy. In other words, the use of robots may weaken human-care relationships. Related to this one of participants expressed:

"I imagine that the worst thing that could happen after using a robot in therapies with my patients would be to lose the emotional bond that therapy normally generates between them and us. It is as if the humanity of the therapy was ignored."

The participants were also asked about the robot's role, for which we used context mapping and prioritization domain activities. When participants were asked about a "magic toolii" to help them during their work and care, they answered that they wanted a tool to interpret the child's thinking, emotions, and intentions. Thus, the central role for the robot would be like a “magic wandii” that can read the child’s mind. These findings are also echoed in some of responses returned in the mailbox during the first phase of our work. However, when this topic was discussed in the plenary session, the participants agreed that the focus should be on improving the children’s communication skills, and that this is to be preferred over interpreting the child’s unexpressed beliefs and thoughts. Thus participants affirmed:

“What I would like the most is that my daughter, when she arrives at home after school, could express in some way how she did or how she felt. I think the robot could help her with that.”

In addition, in all discussions about the robot’s role, the participants imagined the robot in the role of mediator and facilitator, i.e., as a natural extension of the familiar intervention approach. The robot was never exclusively imagined or discussed as a therapist. Thus, when the focus group moderator led the discussion to this topic, the discussion quickly turned into the child’s skills, in which the therapist could use the robot as a catalyst to help the child improve their social skills. In this sense, parents and therapists identified that (i) verbal communication, (ii) expressing of emotions and feelings and (iii) functioning in daily life are the main skills that can be worked on in robot assisted therapy. This fact was consistent with the results of the prioritization domain activity, which were summarized in Figures 30 and 31. There, parents and therapists expressed that strengthen skills relevant to dealing with daily challenges, such as personal care, eating, emotional well-being and verbal communication, are more important than other tasks.

Finally, the participant also correlated the robot’s role with the impact that therapists and parents can generate using the robot, which means that the success of the robot-assisted intervention is influenced by the stakeholder’s abilities to handle the robot’s capabilities in each specific situation. In other words, the human factor and the emotional and physical bond between the child and the caregiver are essential features to assure the success of the CRI. The aforementioned fact was especially observed in the robot demonstration sessions in the sensitization phase, where the music-therapist using an unstructured activity and without the robot’s command training was able to conduct the interaction, enriching the therapy. When the demonstration sessions ended, the music-therapist expressed, *“some of the activities developed came up naturally!”* referring to the proprioception, imitation and dancing activities.

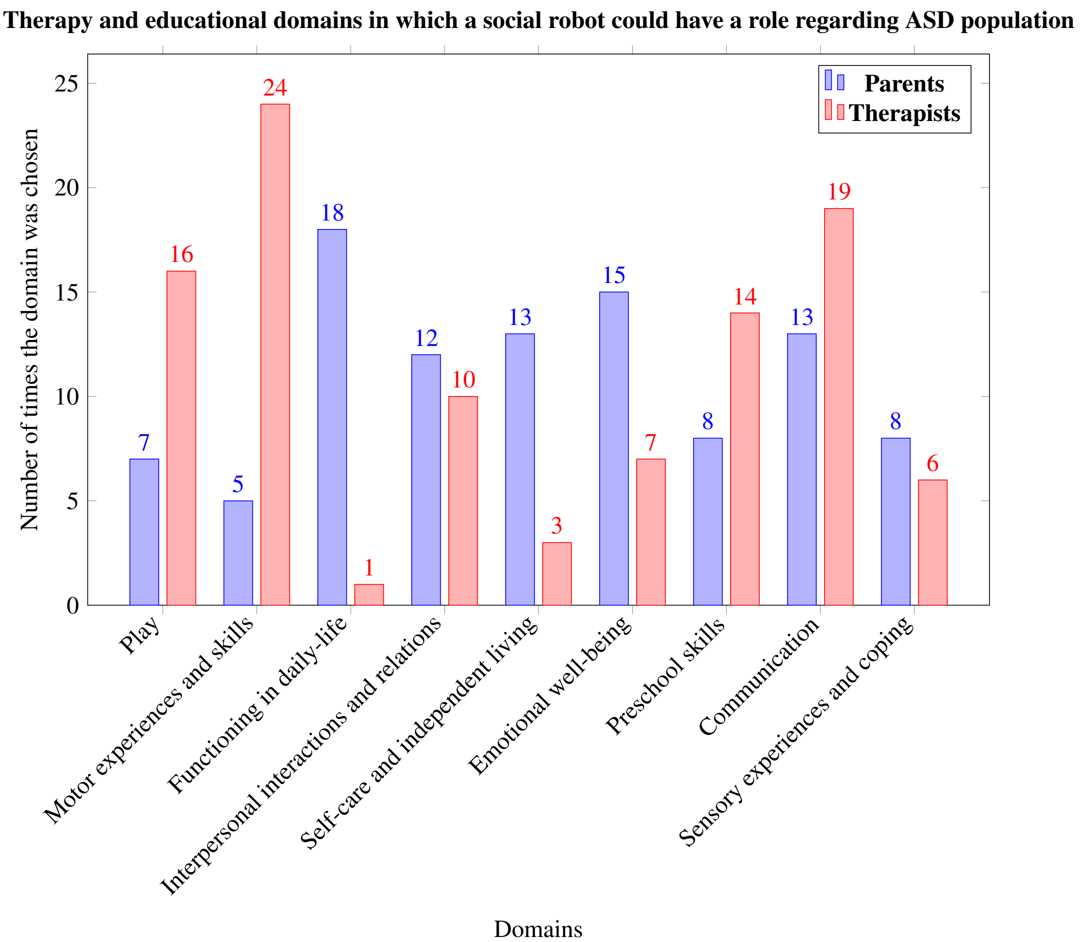

Figure 30 – Results of the prioritization task in the parents and therapist focus groups regarding the main domains for children with ASD.

Additionally, the versatility of the therapist and his/her knowledge about the child's condition can be complemented with a modular and reconfigurable robotic device. In the latest design stage, the community agreed with finding in (RICHARDSON *et al.*, 2018) where authors affirmed: *"No One Autism for All (Nor One Robot for All)"*. Through this study, the allied community wishes to highlight the importance of recognizing ASD for what it is, a diverse set of conditions.

6.3.3 Insights from activities with CwASD

The activities developed with CwASD were challenging but enriching. On the one hand, all the children responded satisfactorily to all phases in part due to the use of material adapted to their needs. However, it is still necessary to make more efforts to encourage children to exhibit behaviors that describe their preferences in greater detail.

Despite the small sample size, the findings indicate that children preferred the

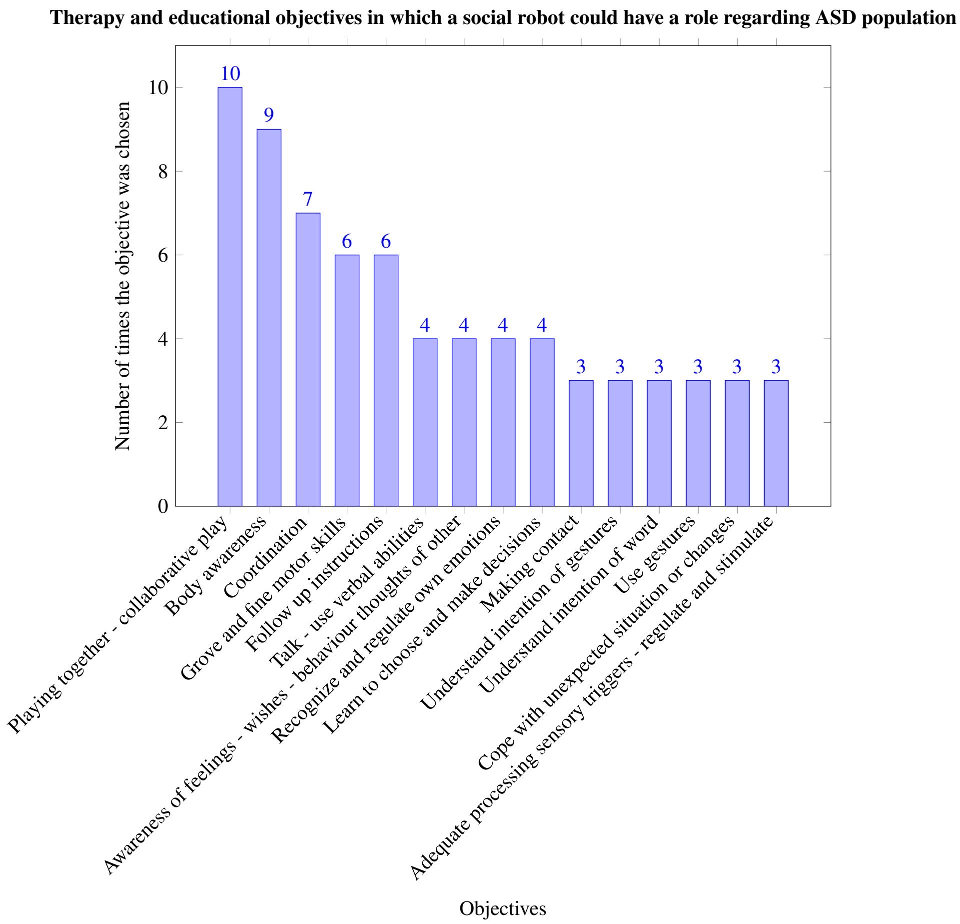

Figure 31 – Results of the therapists' prioritization task regarding the main objectives for children with ASD.

dinosaur Pleo. Perhaps unexpected, Nao and Kaspar were the worst-ranked robots. Trying to explain these findings is difficult, however, based on a qualitative analysis of the children's activities, CASTOR members found that the children showed great enthusiasm when the therapist referred to animals they did not know, such as dragons and dinosaurs. The previous result also can be due to the baby-like features of Pleo, which would be consistent with the stakeholder' view, who affirmed that the robot should look like a baby, in order to allow the CwASD to identify the robot as a peer:

"The robot should have a friendly appearance, has to look kind and look like a child's peer."

The PD process was designed not to impose limits on the ideation by the participants, from which CASTOR distilled guidelines for the design and use of the robot. CASTOR members believe that having lots of ideas and perspectives from different contributors and from different contexts improves the design and increases the positive impact of the design on the community.

The main guidelines of the CASTOR project gathered during the first year are described in Table 9. In this table the community contributions were clustered into four groups, (i) physical requirements, (ii) mechanical and manufacturing features, (iii) technical features and, (iv) intervention implementation. From a general perspective, the described guidelines are consistent with previous requirements regarding appearance and behavior reported in (HUIJNEN *et al.*, 2017), and also provided in (CABIBIHAN *et al.*, 2013). In addition, our generative approach confirms previous evidence described in (PECA *et al.*, 2014) regarding the child's perception and appearance preferences, such as the robot should be visually engaging, have clear facial features and cartoon-like features. Nevertheless, the results presented in this work differ partially in some findings described by they. For example, in this study CwASD preferred Pleo instead of the Probo robot. On the other hand, Romibo was the second choice, and Kaspar and Nao obtained the lowest scores, contrary to reported in (PECA *et al.*, 2014). Even though the sample in this work was smaller, the creative work activity confirms these findings. Also, the generative method directly inquired about the children's preferences through hands-on exercises, providing added value to the results. CASTOR members believe that CwASD in our sample prefer robots that look like a fantastic character, with a preference for a neotenous appearance and with exaggerated features. In addition, a preference was expressed for a robot that is has an appearance sitting between a cartoon and a fantastic animal.

In addition, other important aspects regarding robot's specifications were established, such as, size and proportion, flexibility and modularity. These requirements could be useful to address different specialties of autism therapies and the natural diversity of CwASD.

		Guidelines summary of robot design
Physical Requirements	Aesthetic features	Appearance of a fantastic animal or character (dragon, unicorns, sirens)
		Friendly appearance, look like a baby-featured character
		Neutral gender suitable to be customized
		It can have fantastic elements like tail, chins, crest, wings, etc.
	Body features	With active upper limbs and passive lower limbs (optional active lower limbs)
		Robot proportion around to 2-3 heads with a height between $40-50cm$
		Exaggerated facial features (mouth, two eyes, eyebrows, nose, two ears)
Mechanical and manufacturing features	Modularity	The robotic platform can have interchangeable and adjustable elements, such as nose, ears, hair, etc.
		The robotic platform can have accessories toolkit to customize the interaction, such as clothes, musical instruments, educational tools, and toys
		The robotic platform can have a soft-based structure and appropriate actuators to make a huggable robot
	Materials	Soft materials and different textures, such as silicone, textiles, plushes, leathers, and polymers
		Materials composed of primary colors and without prints or images
Technical features	Voice and sound	It can be equipped with gentle, natural voices (female and male) as well as familiar sounds of animals and musical instruments
		It would be suitable to playback music and video files
	Sensors and actuators	The actuator movements should be gradual, smooth and predictable; should be noiseless and should complement the soft structure
		It can be equipped with vision and touch sensors, microphone and an optional touch screen to produce different stimuli
Intervention implementation	Behaviours and actions	Head and upper limbs movement, look at, point towards, speech, facial expression, eye blinking, reward, grasp objects, hug, play sounds
		Not to use too many stimuli at the same time
	Suggested practices	The robotic platform can be used for different therapies and educational objectives: (i) occupational therapy; (ii) speech and language therapy; (iii) physical therapy; and (iv) psychology
		The intervention could be applied to decrease episodes of anxiety, to engage in communication activities, to feel confident in therapy, to increase motivation and to develop proprioception skill and spatial exploration
		It is crucial to establish a personalized robot-intervention plan to avoid conditioning by the robot's presence and improve the positive therapy effects
		Update the planning regularly

Table 9 – Summary of guidelines resulting from the participatory design process


Illustration by STEVE ASBELL; @rainforestgardn.
Award-winning author/illustrator who fights for inclusion in children's books.
Steve Asbell gave us the permission to use this illustration only for academic purposes

CHAPTER 7

CONCLUSIONS

This research presented a multimodal environment for robot-assisted interventions to assist and enhance traditional practices of ASD diagnosis. The designed framework combines a vision system with automated analysis of children's social behaviors in addition to a robotic platform, both developed upon open source projects. This research contributes to the state-of-the-art with an innovative, flexible and scalable architecture capable of automatically register events of JA, social behaviors, such as adult seeking and patterns of visual contact before and after a robot-based mediation, as well as patterns of behavior related to comfort or discomfort along the ASD intervention. A detailed analysis of the results presented in Chapter 4, as well as, some reflections and conclusions are described below.

7.1 Main findings

From an autism diagnostic perspective, the results obtained in this research are promising, given that they allow exhibiting a series of useful findings to identify autism risk factors associated mainly with the CwASD ability to manifest JA behaviors when a robot elicits them. It is clear that ASD manifestation signs are vast, and our goal was not to find all these patterns. However, in the perspective of diagnosis, this research shows some common signs in CwASD that become more evident when interacting triadically with a robotic platform and a human therapist.

Thanks to the analysis of results at Howard Gardner clinic in Colombia, it was possible to identify that 17 out of 23 children of the CwASD group showed a different behavioral pattern simultaneously in three autism signs: (i) Poor performance in JA tasks in robot-mediated intervention ; (ii) Higher visual preference towards the robot compared with visual contact towards the therapist, as well as a greater interest in physically exploring the robot rather than following its instructions; (iii) The most relevant sign, a low exhibition of adult seeking, i.e., less initiation of protodeclarative and protoimperative behaviors toward therapist/parent to manifest their interest in the robot and to obtain approval to interact. The CwDC frequently showed this last behavior.

It is difficult to compare and validate the results with other studies reported in the literature, given the diagnostic approach used in this study and also because the control group was not formed with TD children. However, using studies from Chapter 2, some coincidences are described, as follows.

The findings of this research concerning the JA score are in agreement with that presented in (ANZALONE *et al.*, 2018), where authors found that the CwASD scored less than the TD children in the RMI. Also, they showed that the CwASD manifested an energy consumption similar to that exhibited by TD children in the RMI. The JA score analysis also is according to the study developed in (DAVID *et al.*, 2018), where they affirmed that the use of RMI does not reflect an improvement in the performance of CwASD for JA tasks compared with the results obtained in TMI. Regarding the behavior of the eye contact, the results of this work are in line with the results in (YUN *et al.*, 2017), where authors showed that the CwASD maintain more eye contact toward the robot than toward the therapist (they attributed this fact due to their preferences for the robotic prompts than those of humans). Additionally, they also found that the mediator effects on the behavioral measurements did not represent an advantage for CwASD. In other works, such as (SCASSELLATI *et al.*, 2018) and (KUMAZAKI *et al.*, 2018a) the authors found improvements in the JA performance of the CwASD using a multi-session approach. Given these variations in the methodological designs among studies, it was not possible to draw a complete comparison of the results.

In this thesis, we showed that, in general, the CwASD did not perform better in the RMI. However, we want to clarify that this does not mean that robots are not suitable to interact with these children as an alternative therapy or diagnostic tool. On the contrary, from the perspective of diagnosis, the use of the robot was relevant, given that it led to more clearly demonstration of behavioral differences between the two groups of children. In fact, this differential behaviors can be interpreted as identification of risk factors that would not have been evident in traditional interventions where CwASD generally exhibit less motivation and little interest in dyadic interaction with the therapist. Through RMI, it was possible to provide systematic implicit and explicit social prompts to open a unique window in which quality behavioral metrics were analyzed.

In a technical perspective, through the MERI system, it was possible to perform face detection, recognition and tracking, landmark detection and tracking, head pose and gaze estimation using a multi-camera approach, and state of the art neural models. This allowed to automatically calculate some metrics related to children's performance, reaching considerable levels of agreement when compared with the results of traditional analyzes. This fact is very promising, considering that MERI system was adapted to three rooms with different dimensions and lighting conditions. Also, it was not necessary to change any structural element of the room, and the cameras were placed in various positions in all cases. Furthermore, the feedback information about the child's performance was successfully used to modulate the supervised behavior of the robot ONO, improving the performance of the CRI and the visual attention of the children. Regarding the VFOA estimation, the algorithm was able to estimate the target into field of view in different situations recurrently. Also, the robot was able to react according to the estimation. However, the algorithm failed when occlusion by the child's hands was generated. On the other hand, the occlusion of therapist and robot was compensated using the multi-camera approach. Thus, the child's face recognition system showed to be imperative to analyze child's behavior in the clinical setup implemented in this research, which required the caregiver's attention in the room. As contribution to the state of art, in this research, a robot-assisted diagnosis technique based on a MERI system was proposed and tested. Through the MERI user interface, JA stimuli were provided, recorded, and analyzed successfully using a well-structured protocol. While social robotic applications are considered a potential tool to elicit differential behaviors, few studies have shown a technological tool to cover all robot intervention stages using experimental designs relevant to core JA challenges.

In conclusion, CwASD demonstrated worse JA compared with a control group during their interaction with the robot, and lower visual contact toward the therapist. However, these children exhibited high preference to look toward the robot and manifested few or no events related to adult seeking behaviour. These three responses were significantly different from those shown by the children in the control group. We suggest that this differential behavior pattern can be used to identify autism risk factors through robot-based interventions. In fact, these findings may represent a meaningful contribution to the literature, take into account the impact of our multimodal environment for robot interventions on autism screening, which provides additional details regarding the suitability of JA protocols elicited by robot mediator for a diagnostic purpose.

Nevertheless, our research has limitations that motivate future research. For instance, the MERI system does not include verbal communication analysis, which remains being a challenge. Also, before the intervention, the participants were not screened, and it was not possible to have the diagnostic measures of the participants due to the internal policies of the HG clinic. In addition, the gender effect was not analyzed, which may be interesting, due to the gender imbalance of the sample, and, in both groups (CwASD and CwDC), the age range of the children was longer than expected.

Finally, it is worth mentioning that the participatory design (PD) process proposed here provided an opportunity to learn from several community actors in the same time, allowing to build guidelines to develop a suitable robotic-device to be implemented for robot-assisted intervention for CwASD. Our PD involved children, caregivers, parents and therapist, each with different cultural and social aspects, who offered insights and scenarios traditionally not considered in robot designs. This broad involvement enriched our PD process and offered an authentic and novel contribution to our research into SAR. Thus, current literature findings regarding methodological and technical requirements of robotic platform and robotic-based interventions for CwASD were complemented through the contributions of our research.

7.2 Contributions of this Ph.D. Thesis

The main contribution of this Ph.D. thesis was the development of a multimodal environment for robot-mediated intervention to assess autism risk factors associated with JA behaviors, as well as, a clinical relevance analysis of the materials and techniques used. Besides, the research and design procedures conducted, reaffirmed in an ethically responsible sense, literature findings regarding the use of RAT and RAD to strengthen ASD therapies and diagnosis. In particular, the research here presented can open an additional chapter concerning CRI's perspectives. This time, regarding how this technique might be used to empower and leverage the caregivers' skills and directly benefit the performance of autism therapies and diagnosis. Finally, there are some important strengths to mention.

1. This work provided, as open source project, access to the developed software and the improvements on Ono's hardware to encourage new researches in this area.

2. Robot mediated interventions were based on evidence-based practices, and clear clinical and research protocols were executed.
3. The outcomes in this Ph.D. thesis showed that CRI scenarios could contribute to *pave the way* to pervasive ASD diagnosis.

7.3 Personal reflections

I would like to express "in a more personal way" my experiences learned as well as my feelings, opinions, and perspectives obtained throughout the development of this research. I want to emphasize that I already completed around 110 h with interventions therapies with children with special needs. The most important feeling that invades me after these hours of interaction with these children is that everything I did (in both aspects, methodological and technical) was worthwhile and was rewarded with every child's smile, progress, or unexpected behavior. At the same time, I have an adverse feeling when I question myself about all the technical aspects that could have been improved to facilitate more the child interaction with the robot in that or another scenario. Therefore, my first reflection is the following; in my opinion, this kind of research only makes sense if it is implemented with the target population (which is the case of this research), regardless of the size of the sample. This reflection is not novel, of course, but sometimes, due to the concern of the engineering, we just put it aside; that's why I wanted to highlight it.

My second reflection continues in a very similar way to the previous one, but this time, it is specific to the robot interaction. At the end of each intervention, independent of the results, I asked myself: what therapeutic benefit attributable to the robot could this intervention represent for the child? The answer is not yet clear to me. However, I want to believe that the fact of them to participate in the intervention and interact with the robot already represents a benefit to them. Although, I am aware that this can't be attributed to the robot and that, on the contrary, it is entirely dependent on the ability of the therapist to get the most out of the robot which triggers a positive impact.

Thus, it is difficult to explain and justify quantitatively and statistically, given that, such as shown in the results, the advantage of using robots in therapy is not reflected in an increase in the performance of CwASD in the sessions, especially in JA intervention. However, for all the people who, like me, participated as observers in many interventions, the positive impact of the interaction with the robot was evident, which was reflected mainly in some signs, such as body posture, interest, and motivation of the child in the intervention and even after the session. But, again for me, it is not that behavioral change that generates therapeutic benefits for the child, but what the therapist can do with the robot at those therapies, for example, to improve some social or communication skills.

Therefore, my feeling at the end of this research is that we need to slightly reinterpret the CwASD-robot interaction. Due to what I saw and experienced, now I am of the people who think that the CwASD have a close emotional bond with their therapists, although this is not evident and, for this reason, I believe that the robot can be used to strengthen this emotional attachment and, thus, achieve more significant therapeutic benefits.

BIBLIOGRAPHY

ANZALONE, S. M.; TILMONT, E.; BOUCENNA, S.; XAVIER, J.; JOUEN, A. L.; BODEAU, N.; MAHARATNA, K.; CHETOUANI, M.; COHEN, D. How children with autism spectrum disorder behave and explore the 4-dimensional (spatial 3D + time) environment during a joint attention induction task with a robot. **Research in Autism Spectrum Disorders**, v. 8, n. 7, p. 814–826, 2014. ISSN 18780237.

ANZALONE, S. M.; XAVIER, J.; BOUCENNA, S.; BILLECI, L.; NARZISI, A.; MURATORI, F.; COHEN, D.; CHETOUANI, M. Quantifying patterns of joint attention during human-robot interactions: An application for autism spectrum disorder assessment. **Pattern Recognition Letters**, Elsevier B.V., v. 0, p. 1–9, 3 2018. ISSN 01678655.

ARESTI-BARTOLOME, N.; GARCIA-ZAPIRAIN, B. Technologies as support tools for persons with autistic spectrum disorder: A systematic review. **International Journal of Environmental Research and Public Health**, v. 11, n. 8, p. 7767–7802, 2014. ISSN 16604601.

ASSOCIATION, A. P. DSM-5 Diagnostic Classification. In: **Diagnostic and Statistical Manual of Mental Disorders**. [S.l.]: American Psychiatric Association, 2013.

BA, S. O.; ODOBEZ, J.-M. Multi-Person Visual Focus of Attention from Head Pose and Meeting Contextual Cues. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 33, n. August, p. 1–16, 2008.

BALTRUŠAITIS, T.; ROBINSON, P.; MORENCY, L. P. 3D Constrained Local Model for rigid and non-rigid facial tracking. **Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition**, p. 2610–2617, 2012. ISSN 10636919.

BALTRUŠAITIS, T.; ROBINSON, P.; MORENCY, L. P. Constrained local neural fields for robust facial landmark detection in the wild. **Proceedings of the IEEE International Conference on Computer Vision**, p. 354–361, 2013.

BALTRUŠAITIS, T.; ROBINSON, P.; MORENCY, L.-P. OpenFace: an open source facial behavior analysis toolkit. **IEEE Winter Conference on Applications of Computer Vision**, 2016. Disponível em: <https://www.cl.cam.ac.uk/~tb346/res/openface.html>.

BALTRUSAITIS, T.; ZADEH, A.; LIM, Y. C.; MORENCY, L. P. OpenFace 2.0: Facial behavior analysis toolkit. **Proceedings - 13th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2018**, p. 59–66, 2018.

BARON-COHEN, S.; ALLEN, J.; GILLBERG, C. Can Autism be Detected at 18 Months?: The Needle, the Haystack, and the CHAT. **British Journal of Psychiatry**, Cambridge University Press, v. 161, n. 6, p. 839–843, 1992.

BARTOLI, L.; GARZOTTO, F.; GELSOMINI, M.; OLIVETO, L.; VALORIANI, M. Designing and evaluating touchless playful interaction for ASD children. **Proceedings of the 2014 conference on Interaction design and children - IDC '14**, n. 1, p. 17–26, 2014. Disponível em: <http://dl.acm.org/citation.cfm?id=2593968.2593976>.

BEAN, J. L.; EIGSTI, I.-m. Assessment of joint attention in school-age children and adolescents. **Research in Autism Spectrum Disorders**, Elsevier Ltd, v. 6, n. 4, p. 1304–1310, 2012. ISSN 1750-9467.

BEGUM, M.; SERNA, R. W.; YANCO, H. A. Are Robots Ready to Deliver Autism Interventions? A Comprehensive Review. **International Journal of Social Robotics**, Springer Netherlands, v. 8, n. 2, p. 157–181, 2016. ISSN 18754805.

BEKELE, E.; CRITTENDON, J. A.; SWANSON, A.; SARKAR, N.; WARREN, Z. E. Pilot clinical application of an adaptive robotic system for young children with autism. **Autism : the international journal of research and practice**, v. 18, n. 5, p. 598–608, 2014. ISSN 1461-7005. Disponível em: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3980197&tool=pmcentrez&rendertype=abstract>.

BEKELE, E.; LAHIRI, U.; SWANSON, A. R.; CRITTENDON, J. A.; WARREN, Z. E.; SARKAR, N. A step towards developing adaptive robot-mediated intervention architecture (ARIA) for children with autism. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, v. 21, n. 2, p. 289–299, 2013. ISSN 15344320.

BELHUMEUR, P. N.; JACOBS, D. W.; KRIEGMAN, D. J.; KUMAR, N. Localizing parts of faces using a consensus of exemplars. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 35, n. 12, p. 2930–2940, 2013. ISSN 01628828.

BELPAEME, T.; BAXTER, P. E.; GREEFF, J. de; KENNEDY, J.; READ, R.; LOOIJE, R.; NEERINCX, M.; BARONI, I.; ZELATI, M. C. Child-Robot Interaction: Perspectives and Challenges. In: **5th International Conference, ICSR 2013**. Bristol, UK: Springer International Publishing, 2013. p. 452–459. ISBN 978-3-319-02674-9.

BELPAEME, T.; BAXTER, P. E.; READ, R.; WOOD, R.; CUAYÁHUITL, H.; KIEFER, B.; RACIOPPA, S.; KRUIJFF-KORBAYOVÁ, I.; ATHANASOPOULOS, G.; ENESCU, V.; LOOIJE, R.; NEERINCX, M.; DEMIRIS, Y.; ROS-ESPINOZA, R.; BECK, A.; CAÑAMERO, L.; HIOLLE, A.; LEWIS, M.; BARONI, I.; NALIN, M.; COSI, P.; PACI, G.; TESSER, F.; SOMMAVILLA, G.; HUMBERT, R. Multimodal Child-Robot Interaction: Building Social Bonds. **Journal of Human-Robot Interaction**, v. 1, n. 2, p. 33–53, 2012. ISSN 2163-0364.

BENTON, L.; VASALOU, A.; KHALED, R.; JOHNSON, H.; GOOCH, D. Diversity for design: A Framework for Involving Neurodiverse Children in the Technology Design Process. **Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14**, p. 3747–3756, 2014.

BERTEL, L. B.; RASMUSSEN, D. M.; CHRISTIANSEN, E. Robots for real: Developing a participatory design framework for implementing educational robots in real-world learning environments. **Lecture Notes in Computer Science (including subseries Lecture Notes in**

Artificial Intelligence and Lecture Notes in Bioinformatics), v. 8118 LNCS, n. PART 2, p. 437–444, 2013. ISSN 03029743.

BOUCENNA, S.; NARZISI, A.; TILMONT, E.; MURATORI, F.; PIOGGIA, G.; COHEN, D.; CHETOUANI, M. Interactive Technologies for Autistic Children: A Review. **Cognitive Computation**, v. 6, n. 4, p. 722–740, 2014. ISSN 18669964.

BRETT, D.; WARNELL, F.; MCCONACHIE, H.; PARR, J. R. Factors Affecting Age at ASD Diagnosis in UK: No Evidence that Diagnosis Age has Decreased Between 2004 and 2014. **Journal of Autism and Developmental Disorders**, Springer US, v. 46, n. 6, p. 1974–1984, 6 2016. ISSN 0162-3257.

CABIBIHAN, J. J.; JAVED, H.; ANG, M.; ALJUNIED, S. M. Why Robots? A Survey on the Roles and Benefits of Social Robots in the Therapy of Children with Autism. **International Journal of Social Robotics**, v. 5, n. 4, p. 593–618, 2013. ISSN 18754791.

CAI, H.; FANG, Y.; JU, Z.; COSTESCU, C.; DAVID, D.; BILLING, E.; ZIEMKE, T.; THILL, S.; BELPAEME, T.; VANDERBORGHT, B.; VERNON, D.; RICHARDSON, K.; LIU, H. Sensing-enhanced Therapy System for Assessing Children with Autism Spectrum Disorders: A Feasibility Study. **IEEE Sensors Journal**, v. 1748, n. c, p. 1–1, 2018. ISSN 1530-437X.

CARVALHO, F. A.; TEIXEIRA, M. C. T. V.; BRUNONI, D.; STRAUSS, V. G.; PAULA, C. S. Identificação de Sinais Precoces de Autismo Segundo um Protocolo de Observação Estruturada : um Estudo de Seguimento. **Psico**, v. 45, n. 4, p. 502–512, 2014. ISSN 0103-5371.

CAZZATO, D.; ADAMO, F.; PALESTRA, G.; CRIFACI, G.; PENNISI, P.; PIOGGIA, G.; RUTA, L.; LEO, M.; DISTANTE, C. Non-intrusive and calibration free visual exploration analysis in children with autism spectrum disorder. **Computational Vision and Medical Image Processing V - Proceedings of 5th Eccomas Thematic Conference on Computational Vision and Medical Image Processing, VipIMAGE 2015**, p. 201–208, 2016.

CHARMAN, T. Why is joint attention a pivotal skill in autism? **Philosophical transactions of the Royal Society of London. Series B, Biological sciences**, v. 358, n. 1430, p. 315–24, 2003. ISSN 0962-8436. Disponível em: <http://rstb.royalsocietypublishing.org/content/358/1430/315.short>.

CHEVALIER, P.; MARTIN, J. C.; ISABLEU, B.; BAZILE, C.; IACOB, D. O.; TAPUS, A. Joint Attention using Human-Robot Interaction: Impact of sensory preferences of children with autism. **25th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2016**, p. 849–854, 2016.

CHONG, E.; CHANDA, K.; YE, Z.; SOUTHERLAND, A.; RUIZ, N.; JONES, R. M.; ROZGA, A.; REHG, J. M. Detecting Gaze Towards Eyes in Natural Social Interactions and Its Use in Child Assessment. **Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.**, v. 1, n. 3, p. 43:1–43:20, 2017. ISSN 2474-9567.

COCO, M. D.; LEO, M.; CARCAGNI, P.; FAMA, F.; SPADARO, L.; RUTA, L.; PIOGGIA, G.; DISTANTE, C. Study of Mechanisms of Social Interaction Stimulation in Autism Spectrum Disorder by Assisted Humanoid Robot. **IEEE Transactions on Cognitive and Developmental Systems**, v. 8920, n. c, p. 1–1, 2017. ISSN 2379-8920.

COECKELBERGH, M.; POP, C.; SIMUT, R.; PECA, A.; PINTEA, S.; DAVID, D.; VANDERBORGHT, B. A Survey of Expectations About the Role of Robots in Robot-Assisted Therapy for Children with ASD: Ethical Acceptability, Trust, Sociability, Appearance, and Attachment. **Science and Engineering Ethics**, Springer Netherlands, v. 22, n. 1, p. 47–65, 2 2016. ISSN 1353-3452. Disponível em: <http://dx.doi.org/10.1007/s11948-015-9649-xhttp://link.springer.com/10.1007/s11948-015-9649-x>.

COSTESCU, C. A.; VANDERBORGHT, B.; DAVID, D. O. Reversal Learning Task in Children with Autism Spectrum Disorder: A Robot-Based Approach. **Journal of Autism and Developmental Disorders**, Springer US, v. 45, n. 11, p. 3715–3725, 2014. ISSN 15733432.

CRISTINACCE, D.; COOTES, T. F. Feature Detection and Tracking with Constrained Local Models. **Procedings of the British Machine Vision Conference 2006**, p. 1–95, 2006.

DANELLJAN, M.; HÄGER, G.; KHAN, F. S.; FELSBERG, M. Accurate Scale Estimation for Robust Visual Tracking. In: **Proceedings of the British Machine Vision Conference 2014**. [S.l.]: British Machine Vision Association, 2014. p. 1–65. ISBN 1-901725-52-9.

DAUTENHAHN, K. A paradigm shift in artificial intelligence: why social intelligence matters in the design and development of robots with human-like intelligence. **50 Years of Artificial Intelligence**, p. 288–302, 2007. ISSN 978-3-540-77295-8. Disponível em: <http://portal.acm.org/citation.cfm?id=1806147>.

DAVID, D. O.; COSTESCU, C. A.; MATU, S.; SZENTAGOTAI, A.; DOBREAN, A. Developing Joint Attention for Children with Autism in Robot-Enhanced Therapy. **International Journal of Social Robotics**, Springer Netherlands, 1 2018. ISSN 1875-4791.

DAWSON, G.; TOTH, K.; ABBOTT, R.; OSTERLING, J.; MUNSON, J.; ESTES, A.; LIAW, J. Early Social Attention Impairments in Autism: Social Orienting, Joint Attention, and Attention to Distress. **Developmental Psychology**, v. 40, n. 2, p. 271–283, 2004. ISSN 00121649.

DEHKORDI, P. S.; MORADI, H.; MAHMOUDI, M.; POURETEMAD, H. R. The Design, Development, and Deployment of RoboParrot for Screening Autistic Children. **International Journal of Social Robotics**, Springer Netherlands, v. 7, n. 4, p. 513–522, 8 2015. ISSN 1875-4791.

EKMAN, P.; FRIESEN, W. Facial Action Coding System. **Consulting Psychologists Press**, 1978.

ESTEBAN, P. G.; BAXTER, P. E.; BELPAEME, T.; BILLING, E.; CAI, H.; CAO, H.-L.; COECKELBERGH, M.; COSTESCU, C.; DAVID, D.; BEIR, A. D.; FANG, Y.; JU, Z.; KENNEDY, J.; LIU, H.; MAZEL, A.; PANDEY, A.; RICHARDSON, K.; SENFT, E.; THILL, S.; PERRE, G. V. D.; VANDERBORGHT, B.; VERNON, D.; HUI, Y.; ZIEMKE, T. How to Build a Supervised Autonomous System for Robot-Enhanced Therapy for Children with Autism Spectrum Disorder. **Paladyn Journal of Behavioral Robotics**, v. 8, n. 1, p. 18–38, 2017. ISSN 2081-4836.

FENG, Y.; JIA, Q.; WEI, W. A Control Architecture of Robot-Assisted Intervention for Children with Autism Spectrum Disorders. **Journal of Robotics**, v. 2018, p. 12, 2018.

FLETCHER-WATSON, S.; ADAMS, J.; BROOK, K.; CHARMAN, T.; CRANE, L.; CUSACK, J.; LEEKAM, S.; MILTON, D.; PARR, J. R.; PELLICANO, E. Making the future together:

Shaping autism research through meaningful participation. **Autism**, p. 1–11, 8 2018. ISSN 1362-3613.

FRAUENBERGER, C.; GOOD, J.; KEAY-BRIGHT, W. Designing technology for children with special needs: Bridging perspectives through participatory design. **CoDesign**, v. 7, n. 1, p. 1–28, 2011. ISSN 15710882.

GOULART, C. M. **UNOBTRUSIVE TECHNIQUE BASED ON INFRARED THERMAL IMAGING FOR EMOTION RECOGNITION IN CHILDREN WITH ASD ROBOT INTERACTION**. 215 p. Tese (Doutorado) — Federal University of Espirito Santo, 2019.

GRYNSZPAN, O.; WEISS, P. L. T.; PEREZ-DIAZ, F.; GAL, E. Innovative technology-based interventions for autism spectrum disorders: A meta-analysis. **Autism**, v. 18, n. 4, p. 346–361, 2014. ISSN 1362-3613. Disponível em: <http://aut.sagepub.com/cgi/doi/10.1177/1362361313476767>.

GUHA, M. L.; DRUIN, A.; FAILS, J. A. COOPERATIVE INQUIRY REVISED Reflections of the Past and Guidelines for the Future of Intergenerational Co-Design. **Action Research: Models, Methods, and Examples**, p. 269–299, 2014.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep Residual Learning for Image Recognition. In: **2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.]: IEEE, 2016. p. 770–778. ISBN 978-1-4673-8851-1. ISSN 1664-1078.

HENDRIKS, N.; SLEGERS, K.; DUYSBURGH, P. Codesign with people living with cognitive or sensory impairments: a case for method stories and uniqueness. **CoDesign**, v. 11, n. 1, p. 70–82, 2015. ISSN 17453755.

HUERTA, M.; LORD, C. Diagnostic evaluation of autism spectrum disorders. **Pediatr Clin North Am**, v. 59, n. 1, p. 103–11, 2012. ISSN 1557-8240.

HUIJNEN, C. A. G. J.; LEXIS, M. A.; JANSENS, R.; WITTE, L. P. de. How to Implement Robots in Interventions for Children with Autism? A Co-creation Study Involving People with Autism, Parents and Professionals. **Journal of Autism and Developmental Disorders**, Springer US, v. 47, n. 10, p. 3079–3096, 2017. ISSN 15733432.

HUIJNEN, C. A. G. J.; LEXIS, M. A. S.; JANSENS, R.; WITTE, L. P. de. Mapping Robots to Therapy and Educational Objectives for Children with Autism Spectrum Disorder. **Journal of Autism and Developmental Disorders**, Springer US, v. 46, n. 6, p. 2100–2114, 6 2016. ISSN 0162-3257.

HUIJNEN, C. A. G. J.; LEXIS, M. A. S.; WITTE, L. P. de. Matching Robot KASPAR to Autism Spectrum Disorder (ASD) Therapy and Educational Goals. **International Journal of Social Robotics**, v. 8, n. 4, p. 445–455, 2016. ISSN 18754805.

HUZOOREE, G.; KHEDO, K. K.; JOONAS, N. Pervasive mobile healthcare systems for chronic disease monitoring. **Health Informatics Journal**, p. 146045821770425, 2017. ISSN 1460-4582.

JOHNSON, C. P.; MYERS, S. M. Identification and Evaluation of Children With Autism Spectrum Disorders. **Pediatrics**, v. 120, n. 5, p. 1183–1215, 2007. ISSN 0031-4005. Disponível em: <http://pediatrics.aappublications.org/cgi/doi/10.1542/peds.2007-2361>.

JORSTAD, A.; DEMENTHON, D.; WANG, I. J.; BURLINA, P. Distributed consensus on camera pose. **IEEE Transactions on Image Processing**, v. 19, n. 9, p. 2396–2407, 2010. ISSN 10577149.

KIM, E.; PAUL, R.; SHIC, F.; SCASSELLATI, B. Bridging the Research Gap: Making HRI Useful to Individuals with Autism. **Journal of Human-Robot Interaction**, v. 1, n. 1, p. 26–54, 2012. ISSN 21630364.

KIM, E. S.; BERKOVITS, L. D.; BERNIER, E. P.; LEYZBERG, D.; SHIC, F.; PAUL, R.; SCASSELLATI, B. Social robots as embedded reinforcers of social behavior in children with autism. **Journal of Autism and Developmental Disorders**, v. 43, n. 5, p. 1038–1049, 2013. ISSN 01623257.

KING, D. E. Dlib-ml: A Machine Learning Toolkit. **Journal of Machine Learning Research**, v. 10, p. 1755–1758, 2009. ISSN 15324435. Disponível em: <http://jmlr.csail.mit.edu/papers/v10/king09a.html>.

KING, D. E. Max-Margin Object Detection. 1 2015.

KNUTSEN, J.; WOLFE, A.; BURKE, B. L.; HEPBURN, S.; LINDGREN, S.; COURY, D. A Systematic Review of Telemedicine in Autism Spectrum Disorders. **Review Journal of Autism and Developmental Disorders**, Review Journal of Autism and Developmental Disorders, v. 3, n. 4, p. 330–344, 2016. ISSN 21957185.

KUMAZAKI, H.; WARREN, Z. E.; SWANSON, A.; YOSHIKAWA, Y.; MATSUMOTO, Y.; ISHIGURO, H.; SARKAR, N.; MINABE, Y.; KIKUCHI, M. Impressions of Humanness for Android Robot may Represent an Endophenotype for Autism Spectrum Disorders. **Journal of Autism and Developmental Disorders**, Springer US, v. 48, n. 2, p. 632–634, 2 2018. ISSN 0162-3257.

KUMAZAKI, H.; YOSHIKAWA, Y.; YOSHIMURA, Y.; IKEDA, T.; HASEGAWA, C.; SAITO, D. N.; TOMIYAMA, S.; AN, K.-m.; SHIMAYA, J.; ISHIGURO, H.; MATSUMOTO, Y.; MINABE, Y.; KIKUCHI, M. The impact of robotic intervention on joint attention in children with autism spectrum disorders. **Molecular Autism**, v. 9, n. 1, p. 46, 9 2018. ISSN 2040-2392.

LE, V.; BRANDT, J.; LIN, Z.; BOURDEV, L.; HUANG, T. S. Interactive Facial Feature Localization. In: . [S.l.]: Springer, 2012. p. 679–692.

LECAVALIER, L. An evaluation of the Gilliam Autism Rating Scale. **Journal of Autism and Developmental Disorders**, v. 35, n. 6, p. 795–805, 2005. ISSN 01623257.

LEMAIGNAN, S.; GARCIA, F.; JACQ, A.; DILLENBOURG, P. From real-time attention assessment to "with-me-ness" in human-robot interaction. **ACM/IEEE International Conference on Human-Robot Interaction**, v. 2016-April, p. 157–164, 2016. ISSN 21672148.

LEMAIGNAN, S.; WARNIER, M.; SISBOT, E. A.; CLODIC, A.; ALAMI, R. Artificial cognition for social human–robot interaction: An implementation. **Artificial Intelligence**, Elsevier B.V., v. 247, p. 45–69, 2017. ISSN 00043702.

LEO, M.; COCO, M. D.; CARCAGNÌ, P.; DISTANTE, C.; BERNAVA, M.; PIOGGIA, G.; PALESTRA, G. Automatic Emotion Recognition in Robot-Children Interaction for ASD Treatment. **Proceedings of the IEEE International Conference on Computer Vision**, v. 2015-Febru, n. c, p. 537–545, 2015. ISSN 15505499.

MALINVERNI, L.; MORAGUIARD, J.; PADILLO, V.; MAIRENA, M.; HERVÁS, A.; PARES, N. Participatory design strategies to enhance the creative contribution of children with special needs. **Proceedings of the 2014 conference on Interaction design and children - IDC '14**, p. 85–94, 2014.

MATSON, J. L.; RIVET, T. T. Characteristics of challenging behaviours in adults with autistic disorder, PDD-NOS, and intellectual disability. **Journal of Intellectual & Developmental Disability**, v. 33, n. 4, p. 323–329, 12 2008. ISSN 1366-8250.

MERTER, S.; HASIRCI, D. A participatory product design process with children with autism spectrum disorder. **CoDesign**, Taylor & Francis, v. 14, n. 3, p. 170–187, 7 2016. ISSN 1571-0882.

MOGHADAS, M.; MORADI, H. Analyzing Human-Robot Interaction Using Machine Vision for Autism screening. In: **2018 6th RSI International Conference on Robotics and Mechatronics (IcRoM)**. [S.l.]: IEEE, 2018. p. 572–576. ISBN 978-1-7281-0127-9.

MOORE, C.; DUNHAM, P. J. **Joint Attention: Its Origins and Role in Development**. [S.l.]: Lawrence Erlbaum, 1995.

NAZNEEN, N.; ROZGA, A.; SMITH, C. J.; OBERLEITNER, R.; ABOWD, G. D.; ARRIAGA, R. I. A Novel System for Supporting Autism Diagnosis Using Home Videos: Iterative Development and Evaluation of System Design. **JMIR mHealth and uHealth**, v. 3, n. 2, p. e68, 2015. ISSN 2291-5222.

NESS, S. L.; MANYAKOV, N. V.; BANGERTER, A.; LEWIN, D.; JAGANNATHA, S.; BOICE, M.; SKALKIN, A.; DAWSON, G.; JANVIER, Y. M.; GOODWIN, M. S.; HENDREN, R.; LEVENTHAL, B.; SHIC, F.; CIOCCIA, W.; PANDINA, G. JAKE® multimodal data capture system: Insights from an observational study of autism spectrum disorder. **Frontiers in Neuroscience**, v. 11, n. SEP, 2017. ISSN 1662453X.

NUNES, D.; WALTER, C. AAC and Autism in Brazil: A Descriptive Review. **International Journal of Disability, Development and Education**, Routledge, v. 00, n. 00, p. 1–17, 2018. ISSN 1465346X. Disponível em: <https://doi.org/10.1080/1034912X.2018.1515424>.

PALESTRA, G.; VARNI, G.; CHETOUANI, M.; ESPOSITO, F. A multimodal and multilevel system for robotics treatment of autism in children. In: **Proceedings of the International Workshop on Social Learning and Multimodal Interaction for Designing Artificial Agents - DAA '16**. New York, New York, USA: ACM Press, 2016. p. 1–6. ISBN 9781450345606.

PAULA, C. S.; RIBEIRO, S. H.; FOMBONNE, E.; MERCADANTE, M. T. Brief report: Prevalence of pervasive developmental disorder in Brazil: A pilot study. **Journal of Autism and Developmental Disorders**, v. 41, n. 12, p. 1738–1742, 2011. ISSN 01623257.

PECA, A.; COECKELBERGH, M.; SIMUT, R.; COSTESCU, C.; PINTEA, S.; DAVID, D.; VANDERBORGHT, B. Robot Enhanced Therapy for Children with Autism Disorders: Measuring Ethical Acceptability. **IEEE Technology and Society Magazine**, v. 35, n. 2, p. 54–66, 6 2016. ISSN 0278-0097.

PECA, A.; SIMUT, R.; PINTEA, S.; COSTESCU, C.; VANDERBORGHT, B. How do typically developing children and children with autism perceive different social robots? **Computers in Human Behavior**, Elsevier Ltd, v. 41, p. 268–277, 12 2014. ISSN 07475632.

PENNISI, P.; TONACCI, A.; TARTARISCO, G.; BILLECI, L.; RUTA, L.; GANGEMI, S.; PIOGGIA, G. **Autism and social robotics: A systematic review**. 2016. 165–183 p.

POUR, A. G.; TAHERI, A.; ALEMI, M.; MEGHDARI, A. Human–Robot Facial Expression Reciprocal Interaction Platform: Case Studies on Children with Autism. **International Journal of Social Robotics**, Springer Netherlands, v. 10, n. 2, p. 179–198, 2018. ISSN 18754805.

QUIGLEY, M.; GERKEY, B.; CONLEY, K.; FAUST, J.; FOOTE, T.; LEIBS, J.; BERGER, E.; WHEELER, R.; NG, A. ROS : an open-source Robot Operating System. In: **ICRA workshop on open source software**. [S.l.: s.n.], 2009. p. 5.

RAHMANI, A. M.; GIA, T. N.; NEGASH, B.; ANZANPOUR, A.; AZIMI, I.; JIANG, M.; LILJEBERG, P. Exploiting smart e-Health gateways at the edge of healthcare Internet-of-Things: A fog computing approach. **Future Generation Computer Systems**, Elsevier B.V., v. 78, p. 641–658, 2018. ISSN 0167739X.

RAMÍREZ-DUQUE, A. A.; FRIZERA-NETO, A.; BASTOS, T. F. Robot-Assisted Autism Spectrum Disorder Diagnostic Based on Artificial Reasoning. **Journal of Intelligent & Robotic Systems**, 3 2019. ISSN 0921-0296.

REHG, J. M.; ABOWD, G. D.; ROZGA, A.; ROMERO, M.; CLEMENTS, M. A.; SCLAROFF, S.; ESSA, I.; OUSLEY, O. Y.; LI, Y.; KIM, C.; RAO, H.; KIM, J. C.; PRESTI, L. L.; ZHANG, J.; LANTSMAN, D.; BIDWELL, J.; YE, Z. Decoding Children's Social Behavior. **2013 IEEE Conference on Computer Vision and Pattern Recognition**, p. 3414–3421, 2013. ISSN 10636919.

REHG, J. M.; ROZGA, A.; ABOWD, G. D.; GOODWIN, M. S. Behavioral Imaging and Autism. **IEEE Pervasive Computing**, v. 13, n. 2, p. 84–87, 4 2014. ISSN 1536-1268.

RICHARDSON, K.; COECKELBERGH, M.; WAKUNUMA, K.; BILLING, E.; ZIEMKE, T.; GOMEZ, P.; VANDERBORGHT, B.; BELPAEME, T. Robot Enhanced Therapy for Children with Autism (DREAM): A Social Model of Autism. **IEEE Technology and Society Magazine**, v. 37, n. 1, p. 30–39, 3 2018. ISSN 0278-0097.

ROBINS, B.; DAUTENHAHN, K. Tactile Interactions with a Humanoid Robot: Novel Play Scenario Implementations with Children with Autism. **International Journal of Social Robotics**, v. 6, n. 3, p. 397–415, 2014. ISSN 18754805.

ROBINS, D. L.; DUMONT-MATHIEU, T. M. Early screening for autism spectrum disorders: update on the modified checklist for autism in toddlers and other measures. **Journal of developmental and behavioral pediatrics : JDBP**, v. 27, n. 2 Suppl, p. 111–9, 4 2006. ISSN 0196-206X.

Robotics Technology Consortium. **A roadmap for us robotics: From internet to robotics**. [S.l.]: Georgia Institute of Technology, 2013. 137 p.

RUDOVIC, O.; LEE, J.; DAI, M.; SCHULLER, B.; PICARD, R. W. Personalized machine learning for robot perception of affect and engagement in autism therapy. **Science Robotics**, v. 3, n. 19, p. eaao6760, 6 2018. ISSN 2470-9476.

SAADATZI, M. N.; PENNINGTON, R. C.; WELCH, K. C.; GRAHAM, J. H. Small-Group Technology-Assisted Instruction: Virtual Teacher and Robot Peer for Individuals with Autism

Spectrum Disorder. **Journal of Autism and Developmental Disorders**, Springer US, v. 0, n. 0, p. 1–15, 2018. ISSN 15733432.

SALDIEN, J.; GORIS, K.; VANDERBORGHT, B.; VANDERFAEILLIE, J.; LEFEBER, D. Expressing emotions with the social robot probo. **International Journal of Social Robotics**, v. 2, n. 4, p. 377–389, 2010. ISSN 18754791.

SARAGIH, J. M.; LUCEY, S.; COHN, J. F. Deformable model fitting by regularized landmark mean-shift. **International Journal of Computer Vision**, v. 91, n. 2, p. 200–215, 2011. ISSN 09205691.

SARTORATO, F.; PRZYBYLOWSKI, L.; SARKO, D. K. Improving therapeutic outcomes in autism spectrum disorders: Enhancing social communication and sensory processing through the use of interactive robots. **Journal of Psychiatric Research**, Elsevier Ltd, v. 90, p. 1–11, 2017. ISSN 18791379.

SCASSELLATI, B. How Social Robots Will Help Us to Diagnose, Treat, and Understand Autism. **Robotics Research**, p. 552–563, 2007. ISSN 16107438.

SCASSELLATI, B.; BOCCANFUSO, L.; HUANG, C.-M.; MADEMTZI, M.; QIN, M.; SALOMONS, N.; VENTOLA, P.; SHIC, F. Improving social skills in children with ASD using a long-term, in-home social robot. **Science Robotics**, Science Robotics, v. 3, n. 21, p. eaat7544, 2018.

SCASSELLATI, B.; Henny Admoni; MATARIĆ, M. J. Robots for Use in Autism Research. **Annual Review of Biomedical Engineering**, v. 14, n. 1, p. 275–294, 2012. ISSN 1523-9829.

SHEIKHI, S.; ODOBEZ, J.-M. Combining dynamic head pose-gaze mapping with the robot conversational state for attention recognition in human-robot interactions. **Pattern Recognition Letters**, v. 66, p. 81–90, 2015. ISSN 01678655.

SHICK, A. Romibo robot project. In: **ACM SIGGRAPH 2013 Studio Talks on - SIGGRAPH ’13**. New York, New York, USA: ACM Press, 2013. (SIGGRAPH ’13), p. 1–1. ISBN 9781450323437.

SHORT, E. S.; DENG, E. C.; FEIL-SEIFER, D.; MATARIĆ, M. J. Understanding Agency in Interactions Between ChildrenWith Autism and Socially Assistive Robots. **Journal of Human-Robot Interaction**, v. 6, n. 3, p. 21, 12 2017. ISSN 2163-0364.

SIMUT, R.; VANDERFAEILLIE, J.; PECA, A.; PERRE, G. Van de; VANDERBORGHT, B. Children with Autism Spectrum Disorders Make a Fruit Salad with Probo, the Social Robot: An Interaction Study. **Journal of Autism and Developmental Disorders**, Springer US, v. 46, n. 1, p. 113–126, 2016. ISSN 15733432.

SMITH, T. Discrete Trial Training in the Treatment of Autism. **Focus on Autism and Other Developmental Disabilities**, v. 16, n. 2, p. 86–92, 2001.

STEINER, A. M.; GOLDSMITH, T. R.; SNOW, A. V.; CHAWARSKA, K. Practitioner’s Guide to Assessment of Autism Spectrum Disorders in Infants and Toddlers. **Journal of Autism and Developmental Disorders**, v. 42, n. 6, p. 1183–1196, 6 2012. ISSN 0162-3257.

TAPUS, A.; MATARIC, M.; SCASSELLATI, B. Socially assistive robotics [Grand Challenges of Robotics]. **IEEE Robotics & Automation Magazine**, v. 14, n. 1, p. 35–42, 3 2007. ISSN 1070-9932.

THABTAH, F.; KAMALOV, F.; RAJAB, K. A new computational intelligence approach to detect autistic features for autism screening. **International Journal of Medical Informatics**, Elsevier Ireland Ltd, v. 117, p. 112–124, 9 2018. ISSN 13865056.

UNIYAL, D.; RAYCHOUDHURY, V. Pervasive Healthcare: A Comprehensive Survey of Tools and Techniques. **Indian Institute of Technology Roorkee**, p. 1–48, 2014.

VALICENTI-MCDERMOTT, M.; HOTTINGER, K.; SEIJO, R.; SHULMAN, L. Age at diagnosis of autism spectrum disorders. **Journal of Pediatrics**, Elsevier, v. 161, n. 3, p. 554–556, 2012. ISSN 00223476.

VALLÈS-PERIS, N.; ANGULO, C.; DOMÈNECH, M. Children's imaginaries of human-robot interaction in healthcare. **International Journal of Environmental Research and Public Health**, v. 15, n. 5, p. 970, 5 2018. ISSN 16604601.

VANDERBORGHT, B.; SIMUT, R.; SALDIEN, J.; POP, C.; RUSU, A. S.; PINTEA, S.; LEFEBER, D.; DAVID, D. O. Using the social robot probo as a social story telling agent for children with ASD. **Interaction Studies**, v. 13, n. 3, p. 348–372, 2012. ISSN 1572-0373. Disponível em: <http://www.jbe-platform.com/content/journals/10.1075/is.13.3.02van>.

VANDEVELDE, C.; SALDIEN, J.; CIOCCI, C.; VANDERBORGHT, B. The use of Social Robot Ono in Robot Assisted Therapy. **International Conference on Social Robotics, Proceedings**, m, 2013.

VARSHNEY, U. Pervasive healthcare and wireless health monitoring. **Mobile Networks and Applications**, v. 12, n. 2-3, p. 113–127, 2007. ISSN 1383469X.

WAINER, J.; DAUTENHAHN, K.; ROBINS, B.; AMIRABDOLLAHIAN, F. A Pilot Study with a Novel Setup for Collaborative Play of the Humanoid Robot KASPAR with Children with Autism. **International Journal of Social Robotics**, v. 6, n. 1, p. 45–65, 2014. ISSN 18754805.

WARREN, Z. E.; ZHENG, Z.; SWANSON, A. R.; BEKELE, E.; ZHANG, L.; CRITTENDON, J. A.; WEITLAUF, A. F.; SARKAR, N. Can Robotic Interaction Improve Joint Attention Skills? **Journal of Autism and Developmental Disorders**, Springer US, v. 45, n. 11, p. 3726–3734, 2015. ISSN 15733432. Disponível em: <http://dx.doi.org/10.1007/s10803-013-1918-4>.

WASHINGTON, P.; VOSS, C.; KLINE, A.; HABER, N.; DANIELS, J.; FAZEL, A.; DE, T.; FEINSTEIN, C.; WINOGRAD, T.; WALL, D. SuperpowerGlass: A Wearable Aid for the At-Home Therapy of Children with Autism. **Proc. ACM Interact. Mob. Wearable Ubiquitous Technol**, v. 1, n. 22, p. 1–22, 2017. ISSN 2474-9567.

WIGGINS, L. D.; BAIO, J.; RICE, C. Examination of the Time Between First Evaluation and First Autism Spectrum Diagnosis in a Population-based Sample. **Journal of Developmental & Behavioral Pediatrics**, v. 27, n. 2, p. 79–87, 2006.

WIGGINS, L. D.; REYNOLDS, A.; RICE, C. E.; MOODY, E. J.; BERNAL, P.; BLASKEY, L.; ROSENBERG, S. A.; LEE, L.-C.; LEVY, S. E. Using Standardized Diagnostic Instruments to Classify Children with Autism in the Study to Explore Early Development. **Journal of Autism and Developmental Disorders**, Springer US, v. 45, n. 5, p. 1271–1280, 5 2015. ISSN 0162-3257.

YUN, S.-S.; CHOI, J.; PARK, S.-K.; BONG, G.-Y.; YOO, H. Social skills training for children with autism spectrum disorder using a robotic behavioral intervention system. **Autism Research**, v. 10, n. 7, p. 1306–1323, 2017. ISSN 19393792.

ZADEH, A.; LIM, Y. C.; BALTRUSAITIS, T.; MORENCY, L.-P. Convolutional Experts Constrained Local Model for 3D Facial Landmark Detection. In: **2017 IEEE International Conference on Computer Vision Workshops (ICCVW)**. [S.l.]: IEEE, 2017. v. 2018-Janua, p. 2519–2528. ISBN 978-1-5386-1034-3.

ZAQUEU, L. d. C. C.; TEIXEIRA, M. C. T. V.; CARVALHO, F. A.; PAULA, C. S. d. Associações entre Sinais Precoces de Autismo, Atenção Compartilhada e Atrasos no Desenvolvimento Infantil. **Psicologia: Teoria e Pesquisa**, v. 31, n. 3, p. 293–302, 2015. ISSN 1806-3446. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-37722015000300293&lng=pt&nrm=iso&tlng=en>.

ZHENG, Z.; ZHANG, L.; BEKELE, E.; SWANSON, A.; CRITTENDON, J. A.; WARREN, Z. E.; SARKAR, N. Impact of robot-mediated interaction system on joint attention skills for children with autism. **IEEE International Conference on Rehabilitation Robotics**, 2013. ISSN 19457898.

ZHENG, Z.; ZHAO, H.; SWANSON, A. R.; WEITLAUF, A. S.; WARREN, Z. E.; SARKAR, N. Design, Development, and Evaluation of a Noninvasive Autonomous Robot-Mediated Joint Attention Intervention System for Young Children with ASD. **IEEE Transactions on Human-Machine Systems**, v. 48, n. 2, p. 125–135, 2018. ISSN 21682291.

ZUBRYCKI, I.; KOLESINSKI, M.; GRANOSIK, G. A participatory design for enhancing the work environment of therapists of disabled children. In: **25th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2016**. New York, USA: [s.n.], 2016. p. 781–786. ISBN 9781509039296.