Impact des préférences sensorielles chez les individus souffrant de troubles du spectre autistique sur leur interaction sociale avec un robot

Pauline Chevalier

NNT : 2016SACLY017

THESE DE DOCTORAT
DE
L'UNIVERSITE PARIS-SACLAY
PREPAREE A
"L'ENSTA-PARISTECH"

ECOLE DOCTORALE N° (573)
INTERFACE

Spécialité de doctorat (Informatique)

Par

Mme Pauline Chevalier

Impact of Sensory Preferences in Individuals with Autism Spectrum Disorder on their Social Interaction with a Robot

Thèse présentée et soutenue à « Palaiseau », le « 08/12/2016 » :

Composition du Jury :

M, Alami, Rachid	Dir. de Recherche, LAAS/CNRS	Président du jury
Mme, André, Elisabeth	Professeur, Augsburg University	Rapporteur
Mme, Paiva, Ana	Professeur, Lisbon University	Rapporteur
M, Scassellati, Brian	Professeur, Yale University	Examinateur
M, Amori, Michel-Ange	Professeur, Université Paris-Sud	Examinateur
Mme, Tapus, Adriana	Professeur, ENSTA ParisTech	Directeur de thèse
M, Martin, Jean-Claude	Professeur, Univ. Paris-Sud, LIMSI-CNRS	Co-directeur de thèse
M, Isableu, Brice	Professeur, Université Aix-Marseille	Co-directeur de thèse
Mme, Nadel, Jacqueline	DR CNRS, UPMC	Invitée
M, Bazile, Christophe	GAPAS	Invité

I dedicate this PhD Thesis to my beloved cats, Mimi and Zézé, for their love and their help during the writing, the cats of Ecole Polytechnique, among them Witty and Kitty, for bringing warmth in the cold gray winter of Palaiseau, and all the kittens in the world.

Résumé

L'objectif de ce travail de thèse est de permettre sur le long terme de proposer des interactions sociales personnalisées exploitant l'attirance que les personnes souffrant de Troubles du Spectre Autistique (TSA) ont envers les robots humanoïdes, tels que Nao (Softbank Robotics), afin d'améliorer leurs capacités d'interaction sociale. Trois institutions spécialisées pour personnes atteintes de TSA participent à notre étude : l'IME Notre Ecole et l'IME MAIA, Institut médico-éducatif pour enfants et adolescents atteints de TSA et la Lendemaine, Foyer d'Aide Médicalisée pour adultes atteints de TSA.

Les différences inter-individuelles sont très présentes dans les TSA et elles impactent différemment le comportement de chaque individu avec TSA dans sa vie (par exemples la communication, l'attention jointe, ou encore les troubles moteurs, à des degrés différents pour chaque individu), et dans cette étude, durant leur interaction sociale avec un robot.

Afin d'envisager à long terme une interaction personnalisée pour chaque participant, une première étape a consisté à définir leur profil sensorimoteur. L'hypothèse qui guide notre étude est que l'intégration des informations visuelles et proprioceptives (perception, consciente ou non, de la position et des changements des différentes parties du corps) d'une personne joue un rôle sur ses capacités sociales. Une personne qui réagit peu aux informations visuelles et qui utilise les informations proprioceptives relatives à ses mouvements ou à la position de son corps de manière exacerbée, aurait plus de difficultés à s'engager et à maintenir une interaction sociale.

Les profils sensoriels des participants ont été évalués à l'aide du test du Profil Sensoriel de Dunn et du test sensorimoteur impliquant une scène mobile virtuelle afin d'évaluer leur dépen-

dance visuelle et proprioceptive. L'analyse des données a permis de classer nos participants en trois groupes montrant des comportements différents face aux informations proprioceptives et visuelles, et à leur intégration.

Nous avons ensuite étudié les liens entre les profils sensoriels des participants et leurs différents comportements sociaux à travers plusieurs tâches impliquées dans les interactions sociales : (1) reconnaissance d'émotions exprimées par deux robots, un avatar et une personne ; (2) interaction sociale avec le robot Nao sur la salutation ; (3) attention conjointe avec le robot Nao, et (4) imitation avec le robot Nao. Cette dernière tâche a fait l'objet de sessions répétées sur huit semaines (modèle de thérapie sur l'apprentissage et de renforcement de l'imitation pour enfants avec TSA).

A travers ces études, nous avons pu observer que les participants ayant une plus forte dépendance à la proprioception et une indépendance au champ visuel ont eu plus de difficultés à interagir avec le robot (moins de regards vers le robot, moins de réponses à l'attention conjointe, plus de difficultés à reconnaitre les émotions, et à imiter un partenaire) que les autres participants.

Nous avons pu observer que les sessions avec le robot Nao ont eu un effet bénéfique chez les participants avec TSA. A la suite des sessions répétées avec le robot Nao, les participants ont montré une amélioration de leurs capacités sociales (regard vers le partenaire, imitations) vers un partenaire d'imitation humain. Ces résultats confortent l'idée d'utiliser les profils sensoriels des personnes avec TSA pour leur proposer, dans des recherches futures, des interactions personnalisées avec les robots.

Summary

The goal of this thesis is to provide contributions that will help in the long term to enable personalized robot-based social interaction for individuals with Autism Spectrum Disorders (ASD). This work was done in collaboration with three care facilities for people suffering from ASD: IME MAIA (France) and IME Notre Ecole, medical and educative schools for children and teenagers with ASD, and FAM La Lendemaine (France), a medical house for adults with ASD.

Inter-individual differences are present in ASD, and impact the behaviors of each individual in their lives, and in this study, during their interactions with a robot.

The first step of our work was to propose an appropriate method to define the proprioceptive and visual profiles of each of our participants. We based our work on the hypothesis that the proprioceptive (the ability of an individual to determine body segment positions (i.e., joint position sense and to detect limb movements in space) and visual integration of cues of an individual with ASD is an indicator of their social and communication skills. We posit that a mitigated behavioral response (i.e., hyporeactivity) to visual motion and an overreliance on proprioceptive information are linked in individuals with ASD to their difficulties in integrating social cues and engaging in successful social interactions.

We used two methods to define the proprioceptive and visual profile of our participant: a well-known questionnaire on sensory preferences and an experimental setup. With the setup, we were able to observe three different groups of postural behaviors in our participants. Thanks to these individual profiles, we could make assumptions on the behaviors that one can expect from each of our participants during interactions with the robot.

We aimed to assess various social skills of our participants in regards to their profiles. We designed three single case studies: (1) emotion recognition with different embodiments (two robots, a virtual agent and a human); (2) a short greeting social task with the robot Nao; and (3) a game evaluating joint attention response to the robot Nao. We also conducted eight weeks-long sessions with an imitation task with Nao.

Through these studies, we were able to observe that the participants that display an overreliance on proprioceptive cues and a hyporeactivity to visual cues had more difficulties to interact with the robot (less gaze towards the robot, less answers to joint attention initiation behaviors, more difficulties to recognize emotions and to imitate a partner) than the other participants.

We were able to observe that the repeated sessions with the robot Nao were benefic for participants with ASD: after the sessions with the robot Nao, the participants showed an improvement in their social skills (gaze to the partner, imitations).

Defining such individual profiles could provide promising strategies for designing successful and adapted Human-Robot Interaction for individuals with ASD.

Acknowledgments

First of all, I would like to express my sincere gratitude to my thesis supervisors, Prof. Adriana Tapus, Prof. Jean-Claude Martin, and Prof. Brice Isableu, for their guidance during my research. Their support, motivation, and their inspiring suggestions have been precious for the development of this thesis content.

Besides my supervisors, I would like to thank the rest of my thesis committee: Prof. Ana Paiva, Prof. Elisabeth André, Prof. Brian Scasselatti, Prof. Rachid Alami, Prof. Michel-Ange Amori, Prof. Jacqueline Nadel, and Dr. Christophe Bazile.

My thoughts also go to the three care centers I worked with during my PhD. Many thanks to the teams that helped me during the experiments answered my question and spent a lot of their time helping with the experiments and answering all my questions. I would like to thank Mrs Gwennaëlle Lerigoleur, Mrs Danièle Lemercier, Mrs Estelle Tronel, and Mrs Sylvie Vax who helped me a lot during the experiments in the care centers.

I would like to thank the participants and their families. My work would not have been possible without their participation and I am very grateful they accepted.

I spent three full years in the U2IS lab in ENSTA ParisTech, and I thank my colleagues for all the good time spent together, work and non-work related.

Finally, I thank my mother, Mrs Sabine Chevalier, my sister, Mrs Louise Chevalier, and my niece, Mrs Emma Chevalier, my fiancé, Mr Gennaro Raiola and Nounours, for believing in me and all their love and unconditional support during these three years.

Table of Contents

List of Figures

List of Tables

Glossary

AASP	Adolescents/Adults Sensory Profiles
ABA	Applied Behavior Analysis
AD#	Adults participants, with # ranging from 1 to 7
ADI-R	Autism Diagnostic Interview-Revised
ADOS-2	Autism Diagnostic Observation Schedule, Second Edition
ADOS-G	Autism Diagnostic Observation Schedule, Generic
AN	Video of the EMBODI-EMO database expressing anger
ASD	Autistic Spectrum Disorder
AU	Action Units
BEAST	Bodily Expressive Action Stimulus Test [De Gelder and Van den Stock, 2011]
BF	Body and Facial condition of the EMBODI-EMO database
BO	Body Only condition of the EMBODI-EMO database
C1	Stable position with static virtual room
C2	Stable position with rolling virtual room
C3	Tandem Romberg position with rolling virtual room
CH#	Children participants, with # ranging from 1 to 12
CoP	Center of Pressure
EMBODI-EMO	Our video database of expressions of emotions
ESCS	Early Social Communication Scales
ESDM	Early Start Denver Model
FAM	Foyer d'Aide Médicalisée

FE	Video of the EMBODI-EMO database expressing fear
FO	Facial Only condition of the EMBODI-EMO database
Fpo	Frequency power at $0.25Hz$
G1	Group G1 in which the participants showed an overreliance on proprioceptive cues and hyporeactivity to visual cues
G2	Group G2 in which the participants showed an even reliance on visual and proprioceptive cues
G3	Group G3 in which the participants showed a hyporeactivity to proprioceptive cues and an overreliance on visual cues
H#	Session with a human partner
HA	Video of the EMBODI-EMO database expressing happiness
HRI	Human-Robot Interaction
ICC	Intraclass Correlation Coefficient
IME	Institut Médico-Educatif
JA	Joint Attention
MSS	Movement sensory sensitivity: tendency to answer quickly to movement stimuli
PECS	Picture Exchange Communication System
Phase (a)	The child imitates the partner (human/robot)
Phase (b)	The partner (human/robot) imitates the child
R#	Session with a robot partner
RMS	Root Mean Square
RVR	Rooling Virtual Room
SA	Video of the EMBODI-EMO database expressing sadness
SAR	Socially Assistive Robotics
SVR	Swaying Virtual Room
TD	Typically Developed

TEACCH	Treatment and Education of Autistic and Related Communication Handicapped Children
TRJA	Time of Response of Joint Attention
VSA	Visual sensation avoiding: tendency to be overwhelmed or bothered by visual sensory stimuli
VSS	Visual sensory sensitivity: tendency to answer quickly to visual stimuli
WOZ	Wizard of Oz

Chapter 1

Introduction

This PhD thesis contributes to the long term development of personalized social interaction between autistic users and robots so as to enhance their social and communication skills. Inter-individual differences are present in Autism Spectrum Disorder, and impact the behavior of each individual during interactions with a robot. The goal of this PhD is to test if the sensory preferences of invididuals with autism impact social interactions with robots [Haswell et al., 2009]. This work was funded by Région Ile de France, Doctoral Fellowship 2013, HANDICAP theme. It took place from December 2013 to December 2016. The subject of this work is multidisciplinary, between the area of robotics, social robotics, sensory integration of proprioceptive and visual cues, affective computing, and autism. The pilot studies conducted in this thesis concern people with autism but more generally people with neurodevelopmenal disorders (ND). In the remainder of this document we will therefore use the more general terms of ND.

1.1 Autism Spectrum Disorder (ASD)

Autism Spectrum Disorder (ASD) is the name for a group of developmental disorders that are characterized by deficits in communication and social skills and the presence of restricted and repetitive patterns of behaviors, interest or activities, as described in the DSM-V [APA, 2013]. The social impairments are largely described in the ASD literature: Individuals with ASD show impaired skills in responding to social attention [Dawson et al., 1998, Dawson et al., 2004], in the recognition of emotions and expressions [Celani et al., 1999, Rump et al., 2009], in imitation task [Williams et al., 2004, Tomasello, 2009, Carpenter et al., 2005, Nadel and Butterworth, 1999], and in initiating and responding to joint attention behaviors [Charman et al., 1997, Mundy and Newell, 2007, Johnson et al., 2007].

A great inter-individual variability is also present in ASD: the symptoms of ASD fall on a continuum [APA, 2013]. Some individuals can show much more severe symptoms than others. In [Milne, 2011], the authors described the fact that numerous studies used sub-groups within their subject pools with individuals with ASD, and that numerous studies showed larger variations around the data collected in groups composed of individuals with ASD than in control groups. They also added that although many cognitive functions deficits are observed in ASD, there are many studies in ASD literature with examples of non-replication, suggesting that no specific cognitive impairment is consistent and universal in ASD. This inter-variability in ASD leads to a necessity to adapt and to personalize therapy for each individual.

Symptoms of ASD are present in the early developmental period, and can be apparent before the age of three years old. ASD is a life-long condition. Very few studies focused on the life-span evolution of autism, but it appears that in some individuals modest improvement in symptoms is evident from childhood to adolescence and into adulthood [Seltzer et al., 2004]. Improvement is not systematic and seems to be lead by the degree of mental retardation [Seltzer et al., 2004, Shattuck et al., 2007]. Individuals with mental retardation showed more symptoms and improved less over time than those without mental retardation [Shattuck et al., 2007].

Studies show a prevalence rate estimate of ASD that vary between 1 to 2% among school-age children, in USA [Baio, 2012], in UK [Baird et al., 2006, Baron-Cohen et al., 2009], and worldwide [Elsabbagh et al., 2012]. In France, the authors in [Chamak et al., 2011] observed that the diagnosis is made earlier in an individual's life with the years: the authors evaluated that the mean age of diagnosis was of 10 ± 8 years from 1960 to 1990, then of 5 ± 3 years from 1990 to 2005, and more recently of 3 ± 1 years from 2003 to 2005. This indicates a better knowledge and screening of the ASD.

Many tools to diagnose, assess, and quantify ASD exist, in commercial form: for examples the Autism Diagnostic Observation Schedule, Generic (ADOS-G) and Second Edition (ADOS-2) [Lord et al., 2000, Rutter et al., 2012], the Autism Diagnostic Interview-Revised (ADI-R) [Lord et al., 1994], the Social Communication Questionnaire [Rutter et al., 2003], and the Social Responsiveness Scale [Constantino and Gruber, 2007]; or in free or open access form: for examples the Autism-Spectrum Quotient [Baron-Cohen et al., 2001b], the Developmental, Dimensional and Diagnostic Interview [Skuse et al., 2004], the M-CHAT/CHAT [Baron-Cohen et al., 2000], and the Children's Social Behavior Questionnaire [Luteijn et al., 2000]. However, there is few information on how these tools are used worldwide. The authors of [Ashwood et al., 2015] showed that in Europe, their are differences in the tools used to diagnose individual with ASD. They emphasize that a convergence of clinical assessment across ASD research and clinical centers in Europe would facilitate collaboration and comparison between clinical trials of novel medications and psychological interventions. In [García-Primo et al., 2014], the authors seek to provide a overview of the screening procedure of children under four years with ASD in Europe, and the methodological concerns and issues associated. They observed a great variability in the methods used to assess ASD across Europe. In addition, they observed that the screening outcomes can be influenced by several factors, among them the age of diagnostic, the level of functioning and autism severity, the parental non-compliance, and the lack of consistency of screening procedures across studies.

Treatments for ASD are very important to help to maximize the individuals' functional independence and to improve their quality of life by minimizing the core ASD features [Myers et al., 2007]. The main objectives are to facilitate development and learning, to enhance social skills and to reduce maladjusted or repetitive behaviors. The support and education of the families is also a major goal. Therapies for individuals with ASD fall in two major categories: educational and medical interventions. Behavioral approaches imply a work on psycho-motor, cognitive and social skills, to reduce aberrant behaviors in individuals with ASD. Several approaches exist. The Applied Behavior Analysis (ABA) [Foxx, 2008], a treatment in which the positive behaviors are reinforced whereas the negative or aggressive behaviors are ignored or punished. ABA intervention requires around 40 hours by week of work with the child with ASD. The Treatment and Education of Autistic and Related Communication Handicapped Children (TEACCH) [Panerai et al., 2002] program as a global approach, based on the collaboration between the parents and the professionals. The program was specifically design for individuals with ASD. It takes into account the features of the ASD and tries to minimize the children' difficulties by using structured and continuous intervention, environmental adaptations, and alternative communication training. The TEACCH program take also into account that individuals with ASD have a wide variability in their symptoms. The Early Start Denver Model (ESDM) [Dawson et al., 2010] is a behavioral intervention for preschool-aged children with ASD that integrates ABA with relationship-based and developmental approaches. The intervention is provided in the children' natural environment (i.e., at home) by trained therapists and parents.

In addition, in individuals with ASD the cognitive processing has been found impaired in the abstraction of information across multiple stimuli or situations and in the generalization of these on new stimuli or situations [Froehlich et al., 2012]. This leads to a difficulty to learn new skills in ASD.

The World Health Organization and the United Nations have identified ASD as an important public health issue across global mental health services [Organization et al., 2013] as there is a huge need to build knowledge in ASD: psychological interventions known to reduce core symptoms and to improve adaptive skills and functioning are available but very resource intensive, and there is a need to strengthen the research efforts to help the individuals with ASD throughout their life.

1.2 Socially Assistive Robotics for Individual with ASD

Since individuals with ASD usually display a strong interest for computers, information and communication technologies have been used with participants with ASD (see

[Grossard and Grynszpan, 2015] for a review). Examples of technologies that have been considered with autistic users include: Virtual Reality [Jung et al., 2006], gaze tracking [Boraston and Blakemore, 2007], virtual characters [Courgeon et al., 2014] and tangible interaction [Farr et al., 2010] and expressive robot [Dautenhahn et al., 2009a].

Currently, research on Socially Assistive Robotics (SAR) is expanding ([Feil-Seifer and Mataric, 2005, Tapus et al., 2007, Matarić and Scassellati, 2016]. Socially Assistive Robotics aims at helping human users through social interaction, rather than through physical contact. One of the target populations is people suffering from ASD. As saw previously, individuals with ASD have impaired skills in communication, interaction, emotion recognition, joint attention, and imitation [Charman et al., 1997]. Many studies have shown that children with ASD have a great affinity for robots, computers, and mechanical components [Hart, 2005, Grossard and Grynszpan, 2015]. In the field of SAR, robots are used as tools in socialization therapies for children with ASD in order to enhance social engagement, imitation, and joint attention skills [Tapus et al., 2012, Kim et al., 2013, Wainer et al., 2010, Kozima et al., 2005].

Research on robots for therapy for individuals with ASD is rapidly growing and focuses on [Feil-Seifer and Mataric, 2005]: (1) designing sociable robots especially modeled for ASD, (2) designing Human-Robot Interaction (HRI) for individual with ASD, i.e., knowledge, methods, and algorithms for natural, transparent HRI that enables humans and robots to interact effectively and cooperatively, and (3) evaluating these robots in therapies for individual with ASD.

Positive effects of robots on ASD participants have been observed. In [Dautenhahn et al., 2009b], the authors set up dyadic interactions through play between autistic children and a robot or human partner. Each child had four interaction sessions, starting with human partners and alternating between human and robot partners. An increased collaborative behavior was observed with a human partner after a session with the robot. In [Kim et al., 2013], the authors used a robot, human, and computer program as a bridge between a human partner and a child with autism in triadic interactions. Results suggest that children with autism display more interest in human partners during the triadic interaction with the robot. In a therapy designed for children with autism [Kozima et al., 2005], the authors found that children with ASD engaged spontaneously in a dyadic play with the robot Keepon thanks to its simple appearance and predictable response. These features succeeded to put Keepon as the pivot of triadic interaction between the child and an adult. In [Vanderborght et al., 2012], the authors used the robot Probo as a social story telling agent for children with ASD. The authors observed that the therapist accompanying the children with ASD had to prompt less the correct behavior to the children when the reader was the robot Probo in comparison to a human. In light of these encouraging findings, many challenges in SAR for individuals with ASD must be addressed. Because of small subject pools and/or short-term experiments, generalized results in the improved skills are often questionable [Scassellati et al., 2012]. The authors in [Salter et al., 2010] discussed how variability (i.e., the environment: experimental or real-life and the autonomy of the robot and/or the participant) of the HRI setup may bias the results in this area. Considering these limitations, the new challenge of assistive robot therapies is to identify how to reduce the bias in such findings. In particular, the authors in [Thill et al., 2012] proposes a new step in robot-assisted therapy; robotic assisted therapeutic scenarios should develop more substantial levels of autonomy, which would allow the robot to adapt to the individual needs of children over longer periods of time.

Thus, although SAR features encouraging results for users with ASD, the individual differences are seldom considered.

1.3 Research Aims

The long-term goal of our research is to develop a new robot-based social interaction model, in addition to standard therapy, for individuals suffering from ND. Since there are important inter-individual differences among individuals with ND (including related to social skills), their interaction with robots needs to be personalized.

1.3.1 Research Hypothesis

An overreliance on proprioceptive information in autism has been suggested in [Haswell et al., 2009]. *Proprioception* can be defined as the ability of an individual to determine body segment positions (i.e., joint position sense) and to detect limb movements (kinesthesia) in space. It is derived from complex somatosensory signals provided to the brain by multiple muscles [Goodwin et al., 1972, Burke et al., 1976, Roll and Vedel, 1982], joints [Ferrell et al., 1987], and skin receptors [Edin, 2001]. Individuals with ASD show normal to exacerbated integration of proprioceptive cues compared to Typically Developed (TD) individuals [Gowen and Hamilton, 2013]. TD individuals have been repeatedly observed to rely more heavily on vision in various perceptivo-cognitive and sensorimotor tasks, following by a progressive age-related declined of visual dependency [Bagust et al., 2013, Assaiante, 1998, Fluckiger and Baumberger, 2003]. Proprioceptive integration in ASD is studied so as to better understand how the contribution of these cues influences interactive and social capacities. In [Haswell et al., 2009, Izawa et al., 2012], the authors observed that the greater the reliance on proprioception, the more children with ASD exhibit impairments in social functions and imitation. Children with ASD were observed to tend to bias proprioceptive over visual feedback when planning a movement [Haswell et al., 2009, Izawa et al., 2012, Miller et al., 2014]. Moreover, limited visual processing skills lead to difficulties in managing social interactions. Vision is an important component for communication and social behaviors. An impairment in visual processing and the gaze function of individuals with ASD may lead to unusual eye contact, difficulty in following the gaze of others or supporting joint attention, and difficulty in interpreting facial and bodily expressions of emotions [Simmons et al., 2009].

Over the TD population, visual field dependent individuals are considered to be more group-oriented and cooperative and less competitive than field independent individuals [Liu and Chepyator-Thomson, 2009]. Visual field dependent individuals are strongly interested in people and get closer to the person with whom they are interacting, whereas visual field independent individuals appear to be cold and detached and they are socially isolated but have good analytic skills [Saracho, 2003].

As we mentionned shortly in the introduction and that will be explained more deeply in the next Section, our participants are patients with more general Neurodevelopmental disorders, so we will hereafter speak about patients with Neurodevelopmental Disorders (ND). We posit that an individual's reliance on proprioceptive and visual cues will affect the way they interact socially, including with a humanoid robot [Haswell et al., 2009, Liu and Chepyator-Thomson, 2009, Saracho, 2003, Coates et al., 1975]. The hypothesis driving our work is that:

- **H1** A mitigated behavioral response (i.e., hyporeactivity) to visual motion and an overreliance on proprioceptive information are linked in individuals with ND to their difficulties in integrating social cues and engaging in successful social interactions.

To the best of our knowledge, no study has examined the sensorimotor and visual profiles of individuals suffering from ND in order to determine and elaborate individualized and personalized scenarios for Human-Robot social interaction therapy.

Thus, since there is not much knowledge on which to rely on, we decided to select an appropriate set of isolated tasks related to social interaction (described below) in order to collect knowledge about how much this hypothesis stands true in different tasks.

1.3.2 Participants

Our work was done in collaboration with three care facilities for people suffering from ASD: *Institut Médico-Educatif* (IME) MAIA (France) and IME Notre Ecole, associations for children and teenagers with ASD, and *Foyer d'Aide Médicalisée* (FAM) La Lendemaine (France), a medical house for adults with ASD. Informed consent for participation was obtained from the parents or by the participants themselves when able. The experimental protocol was approved by the EA 4532 University local ethics committee.

From the care-centers, we recruited only participants that had a medical diagnostic of ASD. Unfortunately, we did not have access to diagnostic reviews as ADOS-G, ADI-R or other ASD diagnostic assessments. So as, we cannot state that our participants have an ASD diagnosis, and we will refer to them as having a Neurodevelopmental Disorders (ND) [APA, 2013]:

> "The neurodevelopmental disorders are a group of conditions with onset in the developmental period. The disorders typically manifest early in development, often before the child enters grade school, and are characterized by developmental deficits that produce impairments of personal, social, academic, or occupational functioning. The range of developmental deficits varies from very specific limitations of learning or control of executive functions to global impairments of social skills or intelligence. The neurodevelopmental disorders frequently co-occur; for example, individuals with autism spectrum disorder often have intellectual disability (intellectual developmental disorder), and many children with attention-deficit/hyperactivity disorder (ADHD) also have a specific learning disorder. For some disorders, the clinical presentation includes symptoms of excess as well as deficits and delays in achieving expected milestones. For example, autism spectrum disorder is diagnosed only when the characteristic deficits of social communication are accompanied by excessively repetitive behaviors, restricted interests, and insistence on sameness."

In the following of the thesis, our research will be focused on studies on individuals with ASD but we are conscious that our participants come from a larger spectrum.

Our subject pool was composed of twelve children and teenagers with ND (11.7$\pm$2.6 years old) and seven adults with ND (26.8$\pm$7.9 years old) from these three care facilities. Five participants are females. A description of each participant's information is given in Table 1.1. We worked with different age group as ND is a life-long condition. For confidentiality purposes, we encoded participant identities as follows: AD#, with # ranging from 1 to 7 for adults participants and CH#, with # ranging from 1 to 12 for children and teenagers participants. The symbols M and F were used to denote male and female, respectively.

1.3.3 Thesis Scope

In this thesis, we aimed to provide contributions that will help in the long term to enable new personalized robot-based social interaction, in addition to standard therapy, for individuals with

Table 1.1: Participants' Description

ID#	Gender	Age	Comments
CH1	M	11	-
CH2	M	9	Suffers of echolalia (i.e., defined as the unsolicited repetition of vocalizations made by another person)
CH3	M	12	-
CH4	M	13	High level of cognition. Asked to be part of the program to meet the Nao robot.
CH5	F	11	Non-verbal; West Syndrome (uncommon to rare epileptic disorder [Dulac et al., 1994]).
CH6	M	13	-
CH 7	M	8	-
CH8	M	15	-
CH9	F	12	-
CH10	M	10	Low cognition
CH11	M	17	-
CH12	M	9	-
AD1	F	25	Microcephaly
AD2	F	25	Diagnosed with Creatine Transporter Deficiency
AD3	M	44	Asperger Syndrome
AD5	M	27	Suffers of epilepsy
AD4	F	21	Febrile seizure
AD6	M	25	Suffers of echolalia; Low cognition
AD7	M	21	-

ND. We based our work on the hypothesis **H1** that the proprioceptive and visual integration of cues of an individual with ND is an indicator of their social and communication skills. The first step of our work was to propose an appropriate method to define the proprioceptive and visual profiles of each of our participants. Then, we aimed to assess various social skills of out participants in regards to their profiles. We designed single case studies including: emotion recognition on different embodiments, a short greeting social task and a game evaluating joint attention response to a robot. After these observations, we designed an eight weeks-long imitation session with a robot. We aimed to adapt the sessions difficulties to the participants and to observe the impact of the participants' profiles on their imitation skills. We also observed the generalization of the skills towards a human partner.

The rest of the thesis is structured in six chapters as follows:

Chapter 2 presents how we defined the profiles of our participants with ND;

Chapter 3 describes the design and validation of the EMBODI-EMO database, which contains facial and bodily video expressions of emotions across different embodiments; this chapter also explains, the impact of the sensory preferences on emotion recognition with our participants with ND;

Chapter 4 reports the first interactive physical encounter between our participants and the robot Nao;

Chapter 5 presents a matching game involving joint attention skills;

Chapter 6 describes a eight weeks long experimental study on imitation with the robot Nao

Chapter 7 presents the overall conclusion of the studies conducted in the thesis and the future directions that we propose.

Chapter 2

Proprioceptive and Visual Profiles for Human-Robot Interaction for Individuals with ND

In this chapter, we explain how we defined proprioceptive and visual profiles of our participant with ND. We base our work on the hypothesis that there is a link between an excessive integration of proprioceptive over visual information and social communication/cognitive disorders (communication, interactions skills, and emotion recognition). We used two methods to define the proprioceptive and visual profile of our participants: a well-known questionnaire on sensory preferences and an experimental setup. With the setup, we were able to observe three different groups of postural behaviors in our participants. Thanks to these individual profiles, we could make assumptions on the behaviors that one can expect from each of our participants during the interactions with the robot (that are described in the following chapters).

2.1 Introduction

Motor, sensory, and visual processing impairments are present in autism but are rarely taken into account in the ASD diagnostic itself [Simmons et al., 2009, Gowen and Hamilton, 2013].

However, these deficits have an influence on the quality of life of individuals suffering from ASD and on their social development. They started to be officially taken into account in 2013 with the publication of the DSM-V [APA, 2013]. In [Bar-Haim and Bart, 2006], the authors observed the impact of motor impairment of children on their social participation in kindergarten. Their observations revealed that children with motor impairments were more likely to have solitary play and less interaction with peers, and therefore, did not explore their physical and social environment, thus leading to social and emotional difficulties. Moreover, visual impairment leads to difficulties in social behaviors. Vision is an important component of communication and social behaviors. An impairment in visual processing and the gaze function in individuals with ASD may lead to unusual eye contact, difficulty in following the gaze of others or supporting joint attention, to integrate contextual visual cues in real-life scenarios, and difficulty in interpreting facial and bodily expressions of emotions [Simmons et al., 2009]. An overreliance on proprioceptive information in autism is also suggested [Gowen and Hamilton, 2013, Greffou et al., 2012, Kohen-Raz et al., 1992, Gepner et al., 1995, Gepner and Mestre, 2002, Haswell et al., 2009]. Individuals with autism show normal to exacerbated integration of proprioceptive cues compared to TD individuals [Gowen and Hamilton, 2013]. More specifically, in [Greffou et al., 2012] and [Kohen-Raz et al., 1992], the authors observed abnormal postural behavior in autism. They found that individuals with ASD show less age-related postural behaviors and are less stable than TD individuals. Results in [Greffou et al., 2012] suggest that postural hyporeactivity to visual information is present in the tested individuals with autism (individuals suffering from ASD with IQs comparable to those of TD individuals). Furthermore, [Gepner et al., 1995] pointed out that individuals with ASD show very poor postural response to visual motion and have movement perception impairments. These results were also observed in [Gepner and Mestre, 2002]. Proprioceptive integration in ASD has been studied to better understand how the contribution of these cues influences interactive and social capacities. In [Haswell et al., 2009], the authors asked 14 children with ASD and 13 TD children to perform a reaching task with their arm while holding a robotic arm that applied a force constraining their movements. The authors observed that the participants built a stronger than normal association between self-generated motor commands and proprioceptive feedback, confirming an overreliance on proprioceptive cues in individuals with ASD. Furthermore, they observed that the greater the reliance on proprioception, the greater the children with ASD exhibit impairments in social functions and imitation. In the continuation of [Haswell et al., 2009], in [Izawa et al., 2012] replicated the same reaching task paradigm, but they provide to the participants visual feedbacks thank to a LED. They found consistent results showing that children with ASD tend to bias proprioceptive over visual feedback when planning a movement. In [Miller et al., 2014], the author observed that in children with ASD, both motor function and visual-motor integration contribute to dyspraxia (i.e., neurodevelopmental disorder affecting performance of skilled gestures, often observed in ASD [Dziuk et al., 2007]) and linked it with the overreliance to proprioceptive cues in ASD. These findings are consistent with the anatomical miswiring in the ASD brain, i.e., the observed overgrowth of short-range white matter connections between somatosensory regions and motor brain regions (these short-range connections favor the movement coding in the intrinsic coordinates of joints and muscles) and an underdevelopment of more distant connections between distant brain regions (e.g., visual-motor and premotor/posterior parietal cortex processing engaged during the planning phase of complex action sequences both favor a coding in extrinsic coordinates of the task) [Herbert et al., 2004]. In addition, the authors in [Chang et al., 2014] found significant association of white matter connectivity with social skills and inattention in children with ASD, TD children and children with Sensory Processing Disorders.

In this chapter, we first explain how we experimentally defined participants' proprioceptive and visual profile with respects to the integration of visual inputs. The chapter is structured as follows: Section 2.2 describe the methodologies used to define the participants' profiles. Section 2.3 presents our analysis method. Section 2.4 presents our results. Finally, Section 2.5 presents an overall conclusion for the chapter.

2.2 Defining Sensory Profiles

The first step of our work was to determine how to define participants' perceptivo-cognitive and sensorimotor profiles. We used two complementary sources of information: (1) the perceptivo-cognitive Adolescents/Adults Sensory Profiles questionnaire (AASP) [Brown and Dunn, 2002] and (2) an experimental sensorimotor setup dedicated to assess the individual's reliance on visual over proprioceptive inputs to control postural balance while confronted to a moving virtual visual room.

2.2.1 The Adolescents/Adults Sensory Profiles (AASP) questionnaire

The AASP developed by Dunn [Brown and Dunn, 2002] was completed by all of our participants. We selected this questionnaire because it has been successfully used in ASD [Jones et al., 2009, Ferguson et al., 2005, Crane et al., 2009]. It enabled us to collect participant's movement, visual, touch, and auditory sensory processing preferences described in terms of the quadrants in Dunn's model of sensory processing [Dunn, 1999]: low registration, sensation seeking, sensory sensitivity, and sensation avoiding. These quadrants are formed by the junctions between individual differences in neurological thresholds for stimulation (high-low) and self-regulation strategies (active-passive):

- Low registration (high-passive): tendency to miss or take a long time to respond to stimuli.
- Sensation Seeking (high-active): tendency to try to create additional stimuli or to look for environments that provide sensory stimuli.
- Sensory Sensitivity (low-passive): tendency to answer quickly to stimuli.
- Sensation Avoiding (low-active): tendency to be overwhelmed or bothered by sensory stimuli and to be actively involved to reduce the stimuli in the environment.

It is thus quite relevant for our research hypotheses that we explained in Section 1.3.1. See Appendix A for more information about the AASP.

As most of our participants do not have the cognitive level to fill the questionnaire by themselves, it was filled with the help of their caregivers who know well their habits and response to everyday life sensory stimuli. In the instruction of the questionnaire, it is specified that the questions can be filled (1) by the person himself/herself; (2) with the help of a caregiver or parent and (3) by a caregiver or parent. Indeed, this questionnaire targets also individuals with intellectual deficiencies. We asked the caregivers to fill the questionnaire as we had a direct contact with them, and were able to inform them well about the conditions and forms of the questionnaire. We assessed Movement, Visual, Touch, and Auditory processing using 29 of the 60 items of the AASP. We eliminated the Taste/Smell processing and Activity level and questions, which were not relevant for the purpose of our study or suitable for individual with invasive ASD's behavior.

2.2.2 Experimental set-up

We designed a sensorimotor experimental setup to assess the sensory integration of each participant. The set up has already been used in several studies [Bray et al., 2004, Isableu et al., 2011]. It evaluates the effect of a moving virtual visual room on postural control and the capability of an individual to use proprioceptive inputs provided in dynamics of balance to reduce visual dependency. In an unstable posture, the integration of proprioceptive feedback differs among individuals [Kluzik et al., 2007, Isableu and Vuillerme, 2006, Isableu et al., 2010]. An individual who integrates proprioceptive cues less than other individuals is called visual dependent. For example, while exposed to visual motion in an unstable posture, the body sway of a visual dependent individual follows the visual stimulus and not the proprioceptive stimulus even if proprioceptive cues were available and reliable [Ravaioli et al., 2005]. Participants were asked to stand quietly in various postural conditions of increasing balance difficulty (normal vs. tandem Romberg) in front of a virtual room, static (SVR) or rolling (RVR) at $0.25Hz$ with an inclination of $\pm 10°$. We chose a rolling frequency of $0.25Hz$: virtual room setups frequently use rolling frequency between $0.1Hz$ to $0.5Hz$ [Cunningham et al., 2006, Gepner et al., 1995, Gepner and Mestre, 2002, Greffou et al., 2008].

It has been found that a frequency of $0.2Hz$ produces the strongest, most synchronized body sway, and that frequencies above $0.5Hz$ produce little body sway [Van Asten et al., 1988]. Participants were asked to stand on a force platform in front of the virtual room; static or rolling in 3 conditions (see Figure 2.1):

- C1 - stable position with SVR: the participant stands on the force platform, straight, feet separated by the length of the hips. The virtual room stays still. The recording lasts 30 seconds.
- C2 - stable position with RVR: the participant stands on the force platform, straight with feet separated by the length of the hips. The virtual room has a sinusoidal movement. The recording lasts 50 seconds.
- C3- tandem Romberg position with RVR: the participant stands on the force platform, straight, one foot in front of the other one. The virtual room has a sinusoidal movement. The recording lasts 50 seconds.

The conditions C1 and C2 permit us to assess the postural behavior in a stable posture with the presence or not of the visual stimulus. In condition C3, the unstable posture coupled to the visual stimulus permit us to assess the dependence to the visual field of an individual [Ravaioli et al., 2005]. Thanks to this setup, we are able to compare the postural behavior between different posture and exposure to the visual stimulus. We used this setup as it was non-invasive (e.g., "vibration-induced illusory joint movements" [Roll and Vedel, 1982], "Reach and Grasp Test" [Haswell et al., 2009]) and it required no cognition skills (e.g., "Rod and Frame Test", "Rubber Hand Illusion" [Cascio et al., 2012]). Similar setups were used with children with ASD in [Gepner et al., 1995, Gepner and Mestre, 2002].

The virtual room was created with Blender, a free 3D rendering software, see Figure 2.2. The virtual room was decorated with child toys and furniture so as to create a friendly environment. A toy plane was placed in the line of sight of the individuals in order to help them to focus on the task, and not to be distracted. The virtual room was displayed on a wall with a short focal projector in a dark room. It rolled at $0.25Hz$ with an inclination of $\pm 10°$. For the adult group

Figure 2.1: Experimental setup for adults participants in condition C3

setup, the dimension of the projection was $2.4m$ large x $1.8m$ high and the participants stood at $1.3m$ of the point of observation. For the children group setup, the dimension was $1.75m$ large x $1.30m$ high and the participants stood at $1m$.This permitted us to maintain the angular diameter around $31°$ in horizontal and $41°$ in vertical in both setup. We investigated if the age of our participants had an influence on their posture behavior. Indeed, in TD individuals, children show more dramatic postural reactions to visual sway than adults [Schmuckler, 1997]. However, as shown in [Gepner and Mestre, 2002], children with ASD showed less response to visual stimulus in virtual room experiments in stable position than TD children. We expected that:

- (1) children participants should show more swaying than adults in stable position and without visual stimulus (C1);
- (2) as we work with participants with ASD, the postural reaction to visual sway should not be influenced by age;
- (3) the effect of the RVR should be maximum in the unstable postural condition (C3);
- (4) the effect of (3) should not be larger in adults than in children.

Figure 2.2: Screenshot of the virtual room used in the experiment, design with Blender

2.3 Data Analysis

For all sessions, a force platform (AMTI OR6-5-1000) was used to record the displacement of the Center of Pressure (CoP) with a sampling frequency of $1KHz$. We used a Butterworth filter with a cut-off frequency of $10Hz$ on the recorded data in order to reduce the noise. We computed the Root Mean Square (RMS) of the displacements of the CoP in mediolateral directions as an indicator of an individual's stability. Indeed, the RMS provided us the information about the variability of the displacements of the CoP in space [Chiari et al., 2002]. We computed the Frequency power at $0.25Hz$ (Fpo) so as to evaluate the postural response to the visual stimulus. More an individual is coupled with the visual stimulus, more Fpo is high. We observed the displacements in mediolateral direction as it is the direction of our visual stimulus. RMS and Fpo should be correlated if our participants with ASD follows the RVR movement:

1. If the RMS and the Fpo are correlated then we can expect these coupling capabilities with contextual cues promise higher social interaction capabilities.

2. If the RMS and the Fpo are not correlated (no coupling), then this might suggest that the visual stimulus is integrated as noise, inducing disorientation and instability.

We performed Repeated Measures Analysis on the RMS and the Fpo for the age groups (adults; children) and the conditions (C1; C2; C3). The significance threshold was set to $p < 0.05$. We used Statistica version 13 to perform the analyses. The participants AD6, CH6, CH7,

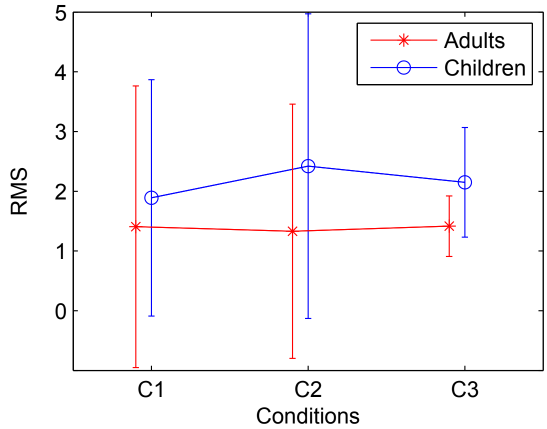


Figure 2.3: Mean RMS for the adults and children in the 3 conditions

and CH12 showed nervousness, and/or agitation during the recording, resulting to dramatic changes in their displacements of the CoP during some recording. Except for the clustering, we excluded them of the analyses.

2.4 Results

2.4.1 Displacements and Root Mean Square (RMS)

As we expected, we found a significant main effect of the participants' age on their RMS ($F(1, 13) = 6.92; p < 0.05$) across all conditions. The mediolateral RMS of the adults was smaller ($M = 0.84$; $SD = 0.57$) than those of the children ($M = 1.70$; $SD = 1.12$), indicating that the children were globally more variable than the adults (see Figure 2.3). The conditions (C1 vs C2 vs C3) did not impact the displacements of the CoP of the participants. The Conditions x Age interaction was not significant on the RMS.

2.4.2 Displacements and Frequency power response at 0.25 Hz (Fpo)

As we expected, the main effect of Conditions was significant on the Fpo ($F(2, 26) = 13.11; p < 0.05$). In the C1 condition ($M = 0.17$; $SD = 0.16$), participants had the smallest Fpo, followed by the C2 condition ($M = 0.34$; $SD = 0.38$) and the C3 condition ($M = 0.67$; $SD = 0.31$) had the highest Fpo. This suggests that the participants' displacements of the

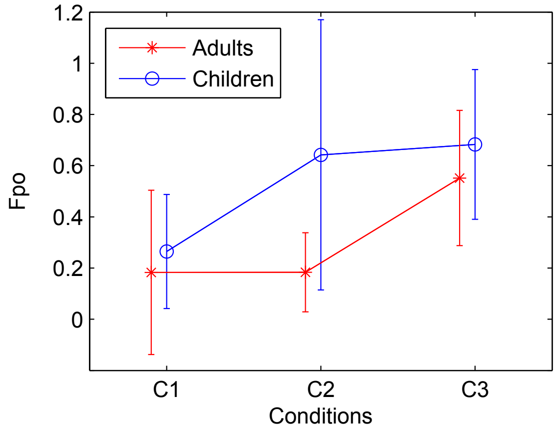


Figure 2.4: Mean Fpo for the adults and children in the 3 conditions

CoP were coupled with the movement of the visual room when they were exposed to it, and that this coupling is maximized in the more difficult stance. In addition, the Fpo was positively correlated with the RMS in the C3 condition ($R = 0.73$; $p < 0.01$), indicating that the larger displacement of the CoP are possibly coupled with the RVR (ie. the visul stimulus). We found a significant main effect of the participants' age ($F(1, 13) = 5.37$; $p < 0.05$). The Fpo of adults was smaller ($M = 0.29$; $SD = 0.33$) than those of children ($M = 0.46$; $SD = 0.36$) across all conditions. This suggests a higher coupling between postural response to the RVR and the displacements of the CoP in children than in adults. However, the Age x Conditions interaction for the Fpo was not significant. In Figure 2.4, we can observe that children and adults postural response to the visual stimulus is similar in the C3 condition, indicating that children with ND do not respond in a higher way to visual cues in comparison to adults with ND. The frequency of $0.25Hz$ is present in natural swaying [Van Asten et al., 1988] and the RMS was higher in children than in adults. As the Age x Conditions interaction for the Fpo was not significant, we can assume that the higher Fpo on the whole experiment was induced by the greater variability of the displacements of the CoP of the children. This result suggests that our participants postural behavior was not driven by age as we expected. We found that our participants' postural coupling to the virtual room was driven by the conditions (C1; C2; C3).

Table 2.1: AASP Items and CoP behavior selected for the Dendrogram Analysis

Movement low registration
Visual low registration
Movement sensation seeking
Visual sensation seeking
Movement sensory sensitivity
Visual sensory sensitivity
Movement sensation avoiding
Visual sensation avoiding
RMS for the C1, C2 and C3 conditions in mediolateral direction
Fpo for the C1, C2 and C3 conditions in mediolateral direction

2.4.3 Grouping of the Participants

We performed a clustering analysis (dendrogram, Ward method) on the AASP items on Movement and Visual sensory preferences, the RMS and the Fpo of all the 19 participants (12 children and seven adults with ND) (see Table 2.1 to see the specific items selected). We sought to identify if the postural response to the visual stimulus and AASP scores were able to discriminate our participants, as we aimed to assess the impact of theses profiles on social personalized interactions with robots.

The dendrogram gave us three groups, see Figure 2.5:

- G1: 8 participants CH3; CH5; CH8; CH10; CH11; AD2; AD4; AD6;
- G2: 7 participants CH1; CH4; CH7; CH9; CH12; AD1; AD3;
- G3: 4 participants CH2; CH6; AD5; AD7.

The RMS in all conditions for the three groups detected are shown in Figure 2.6. Figure 2.7 shows Fpo in the C3 condition. In Figure 2.8, the mean AASP score (Movement sensory sensitivity (MSS), Visual sensory sensitivity (VSS), and Visual sensation avoiding (VSA)) for each group are illustrated.

Repeated Measures Analysis was applied on the RMS and the Fpo for the groups (G1; G2; G3) and the conditions (C1; C2; C3). We found no main effect of the groups on the RMS and on the Fpo of the participants. However, we found significant interaction effect between groups over the conditions on the RMS ($F(4, 24) = 3.55$; $p < 0.05$), see Figure 2.6. Participants from groups G1 and G2 showed great CoP variability in all conditions, unlike participants from group G3 that showed a greater CoP variability only in the C3 condition. Figure 2.6 also informs us that participants in group G1 had their RMS that decrease from the C1 to C2 conditions and from the C2 to C3 conditions, whereas participants in group G3 had their RMS that increased from the C1 to C2 conditions and from the C2 to C3 conditions. This indicates that participants from group G1 maximized the use of proprioception to reduce the effect of the visual stimulus in unstable position. Identically, we found a significant interaction effect between the groups and the conditions on the Fpo ($F(4, 24) = 9.79$; $p < 0.001$), see Figure 2.7. This indicates us that each group has a different postural response toward the visual stimulus. Figure 2.7 suggests that in the C1 condition, participants from groups G1 and G2 showed a higher coupling with the frequency of the visual stimulus than participants from group G3. As in the C1 condition, the participants are not exposed to the RVR, and this result shows the higher instability of these participant in comparison to participants from group G3. The coupling with the rolling visual stimulus is similar in the C2 and C3 conditions for participants from groups G1 and G2, indicating that the difficulty of postural task (stable or unstable) did not increase the strength of the coupling with the rolling virtual room: being in a stable or unstable posture did not affect them in their response to the visual stimulus. Participants from group G3 showed a greater coupling to the visual stimulus in the C3 condition than in other conditions, indicating that they responded more strongly to the visual stimulus in an unstable posture and thus are more visual dependent (see Figure 2.7).

Furthermore, we also examined the correspondences between the scores of different items of the AASP and the data obtained from the CoP recording, listed in Table 2.1:

- High score in MSS (i.e., tendency to answer quickly to movement stimuli) was inversely correlated to Fpo in mediolateral direction for the C3 condition ($R = -0.53$; $p < 0.05$), indicating that participants who revealed by the AASP questionnaire to have a tendency to answer quickly to movement stimuli were less driven by the visual stimulus in unstable position.

- High score in VSS (i.e., tendency to answer quickly to visual stimuli) was positively cor-

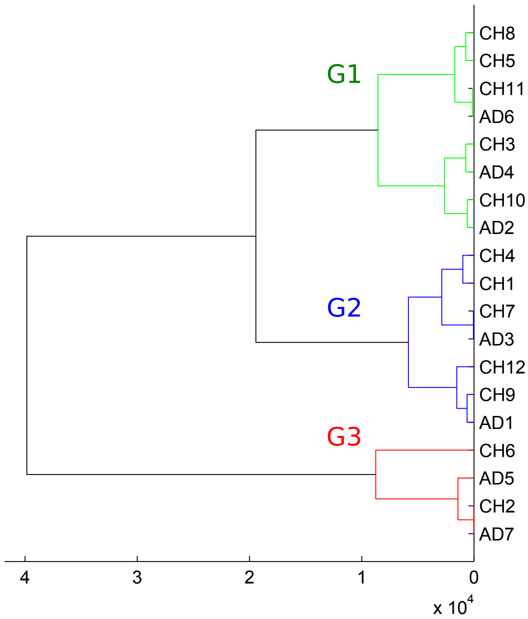


Figure 2.5: Dendrogram Analysis

related to mediolateral RMS for the C3 condition ($R = 0.61$; $p < 0.05$). This suggests that participants who revealed by the AASP questionnaire to have a tendency to answer quickly to visual stimuli were more unstable when exposed to the visual stimulus in unstable position.

- High score in VSA (i.e., tendency to be overwhelmed or bothered by visual sensory stimuli) was positively correlated to mediolateral RMS for the C3 condition ($R = 0.59$; $p < 0.05$), indicating that participants who revealed by the AASP questionnaire to have a tendency to be overwhelmed or bothered by visual stimuli were more unstable when exposed to the visual stimulus in unstable position.

Three of the selected items of the AASP showed to be correlated with the postural response variability to a visual stimulus in an unstable position (C3), confirming that these AASP items match with the behavioral response of the participants.

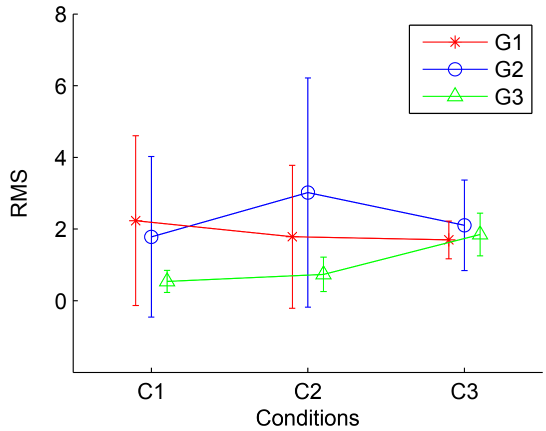


Figure 2.6: Mean Root Mean Square (RMS) for the three groups for the three conditions

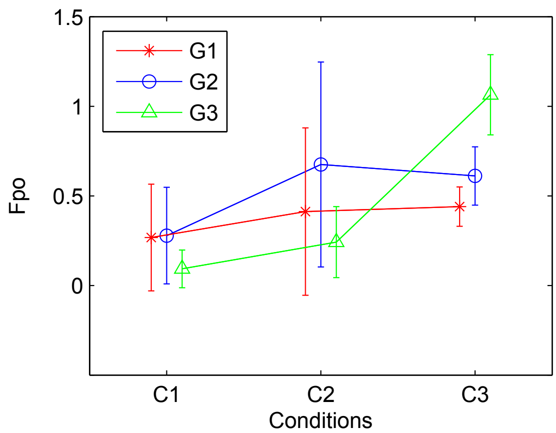


Figure 2.7: Mean Frequency power at $0.25Hz$ (Fpo) for the three groups for the three conditions

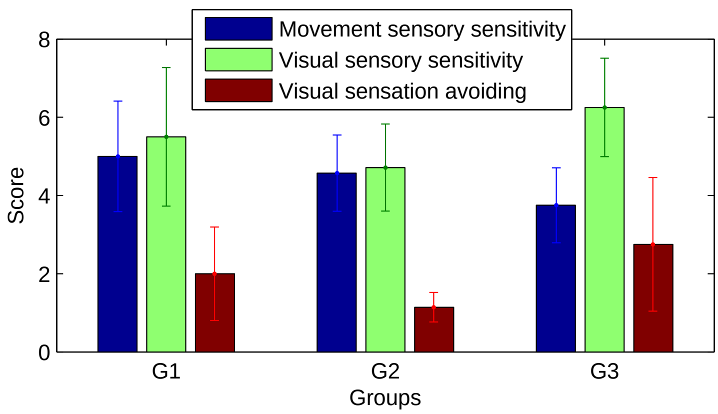


Figure 2.8: Histograms for three relevant Adolescent/Adult Sensory Profile (AASP) items scores of the groups

2.5 Discussion and conclusions

Thanks to the experiment that we described in this Chapter, we succeeded to form three groups by clustering our participants and describing each participant's response to visual and proprioceptive inputs:

- **Group G1** (8 participants: CH3; CH5; CH8; CH10; CH11; AD2; AD4; AD6) includes participants with high scores in MSS, low scores in VSS and low score in VSA. Participants showed strong visual independence to the RVR. This suggests that *these participants have an overreliance on proprioceptive cues* (allowing decreasing their RMS in difficult stance) *and hyporeactivity to visual cues*.

- **Group G2** (7 participants: CH1; CH4; CH7; CH9; CH12; AD1; AD3) includes participants with high scores in MSS, low scores in VSS and low score in VSA. Participants showed moderate reactivity to the RVR. *This suggests that these participants rely evenly on visual and proprioceptive cues*.

- **Group G3** (4 participants: CH2; CH6; AD5; AD7) includes participants with low scores in MSS, high scores in VSS and high score in VSA. Participants showed hypereactivity to the RVR. This suggests that *these participants have a hyporeactivity to proprioceptive cues and an overreliance on visual cues*.

We found that overall, children had greater variability of the displacements of their CoP than adults which was expected (due to biomechanical, anthropometrical, sensory integration factors and maturation of these processes). The different conditions of posture and exposure to the visual stimulus (C1; C2; C3) did not impact the variability of the CoP of our participants. However, the participants' displacements of the CoP showed to be more coupled with the frequency of the visual stimulus in the C3 condition than in the C1 and C2 conditions. We observed that the coupling to the frequency of the visual stimulus was higher in children than in adults on the whole experiment, but not inside the conditions. As the children were experiencing more variability to their CoP and as the frequency $0.25Hz$ (i.e., the frequency of our visual stimulus) is present in natural swaying [Van Asten et al., 1988], we posit that this higher coupling to this frequency on the whole experiment by the children was induced by the greater variability of the displacements of their CoP. This result suggests that our participants postural behavior was not driven by age as we expected.

In TD individuals, children show more dramatic postural reaction to visual stimulus [Schmuckler, 1997]. In children with ASD, it is suggested that sensory integration differs from TD individuals [Gowen and Hamilton, 2013, Haswell et al., 2009, Gepner et al., 1995, Gepner and Mestre, 2002]. We found in this experiment that the age of the participants with ND did not impact the strength of the postural coupling with the visual stimulus, differently to TD individuals. Similarly to TD individuals, children with ND had more variability in the displacement of their CoP in comparison to adults with ND.

We also observed a great variability in our participants' postural behavior. In our results, we found three groups with different sensory integration among our participants. As in [Molloy et al., 2003, Torres et al., 2013] we found in our participants individuals (mixing children and adults) with a weak proprioceptive integration and strong visual dependency. In [Molloy et al., 2003], the authors found an impairment of the proprioception input in autism: children with ASD used more the visual cues to reduce sway and maintain balance. In [Torres et al., 2013], the authors found that unlikely to typically developed individuals; individuals with ASD have an impaired proprioception development. Their sensory-motor signal appears to remain at the kinesthetic stage of typically developed 3–4 years old children, and have to rely on visual inputs. They also conjectured that the impaired proprioception of physical micro-movements of the individuals with ASD impedes as well their visual perception of micro-movements in others during real time interactions, impairing their abilities to interact with people. And, as in [Gepner and Mestre, 2002] and [Haswell et al., 2009], we found in our participants individuals relying more on proprioceptive inputs and weak visual dependency.

With these results and our hypothesis **H1**, we are able to make assumptions on the behaviors that each individual will have during Human-Robot Interaction sessions. Therefore, we posit that *individuals from group G1 will have less successful social interactions than the ones from* *groups G2 and G3, and that individuals from group G3 will have the most successful social interactions.*

Chapter 3

Task #1: Recognition of Expressions of Emotions Displayed by two Robots, a Virtual Agent and a Human

In this chapter, we assess whether the recognition of expressions of emotion in ND is facilitated or reduced by (1) the abundance of nonverbal signals/channels (body and/or face), (2) their embodiment (using 2D or 3D platforms, a virtual agent, a human, and two mini-humanoid robots, respectively) and (3) depends on the proprioceptive and visual cues integration profiles of our participants. Such a database do not exist in the literature. First, we explain how and why we collected EMBODI-EMO and describe the validation of this database with TD individuals. Then, we explain how we used this database to investigate the relationship between recognition of multichannel and embodied expressions of emotion, and proprioceptive and visual cues integration profiles of our 19 participants with ND. The results show that adults with ND have more difficulties to perform the recognition task than children with ND. As the adults' results were close to random, we also performed the analysis only on the children' answers. Our results show that the recognition of emotion is improved when multiple channels are combined but depend on the platform. Visual and proprioceptive cues integration profile of a child with ND influence his/her abilities to recognize emotions. Children with ND relying on proprioceptive information have lower recognition scores than the other children with ND.

3.1 Introduction

Emotion recognition plays an important role in social interactions and has often been used in ASD therapies and experimental studies. Examples include a computer program [Silver and Oakes, 2001], cartoon using animated vehicles with real emotional faces for young children [Golan et al., 2010], expressive robots [Costa et al., 2014a, Pop et al., 2013].

Emotions are expressed through different channels such as facial expressions [Ekman and Friesen, 1984], body movements/postures [Kleinsmith and Bianchi-Berthouze, 2013, Mehrabian, 1972], and speech [Scherer, 1995]. In TD individuals, all the channels interact in emotion recognition [De Gelder et al., 2015]. In [Simmons et al., 2009], the authors observed that people with ASD have difficulties in recognizing emotions expressed through facial and body movements. Moreover, in [Simmons et al., 2009], participants with ASD could not easily describe and recognize emotions from a body in motion and interpret movement in a social context. The face seemed to be interpreted as a complex object without social interest in autistic individuals [Simmons et al., 2009]. The recognition of facial expression of emotions by individuals with autism and TD individuals has been studied using human pictures [Baron-Cohen et al., 1997], virtual agents [Courgeon et al., 2012], and robots [Costa et al., 2014b]. Overall, TD individuals recognized facial expressions better than individuals with ASD. However, in [Begeer et al., 2006], the authors observed that children with ASD had greater difficulty recognizing emotions without context, whereas TD children performed identically, with or without context. Additionally, the authors in [Celani et al., 1999] observed that individuals with ASD have good observation strategies and can indeed detect emotional cues.

In this chapter, we evaluate the capabilities of our participants with autism to recognize expression of emotions via facial expressions, bodily expressions and combinations of both. We compare their recognition of expressions displayed by different embodiments. We want to observe if the channel of expression or the embodiment has an impact on the recognition level and is related to the proprioceptive and visual preferences of an individual with ND. In our work, we intend to propose robot-based interactions, and many robotic-platforms exist. They all propose different features of expression, that may impact the interaction with the participants. We also want to evaluate the impact of virtual agent and a real human to compare the behavior of our participants. Although there are lots or existing databases of videos containing expressions of emotions, we did not find any available database that contains dynamics expressions of facial and bodily expressions of emotions displayed using different embodiment. There are a few databases of bodily static (BEAST [De Gelder and Van den Stock, 2011]) or dynamic (AFFECTME [Kleinsmith et al., 2006], GEMEP [Dael et al., 2012]) expressions of emotions but none of them presents similar bodily expressions with regard to various physical and virtual embodiments. Few research works focus on the development of emotional postures for the Nao robot [Erden, 2013] and for a virtual agent [Coulson, 2004] reaching good recognition levels in TD populations. We thus developed the EMBODI-EMO database[1], a video database of expressions of emotions. We used this database to evaluate individuals' skills to recognize different combination of body and facial expressions of emotions enacted through different embodiments and to analyze relationships between these skills (recognition scores) and a user's proprioceptive and visual profile. In [Atkinson, 2009], adults with ASD were shown to be less accurate than TD adults in classifying emotions and in perceiving movements. The results showed that more the individuals with ASD had difficulties to perceived coherent motion, the less they were accurate on recognizing emotions.

We use and compare different embodiments (i.e., two humanoid robots, one avatar, and one real human), to vary the abundance of signals (bodily and/or face) of expressions of emotions. We use Nao robot from SoftBank Robotics (former Aldebaran Robotics), Zeno (R25) from Hanson Robotics, a female virtual agent (Mary) from the Multimodal Affective and Reactive Character (MARC) platform [Courgeon and Clavel, 2013], and a female human (Human) (Figure 3.1).

[1]Database link: `http://perso.ensta-paristech.fr/~tapus/eng/media/EMBODI-EMO.zip`

Figure 3.1: Different embodiments in the neutral pose, from top left to bottom right: Nao-robot, Zeno-robot, Mary-virtual agent, and Human.

From **H1**, we made the following hypothesis regarding the impact of proprioceptive and visual integration profiles on an individual's emotion recognition performance:

- **H1'** Individuals with an overreliance on proprioceptive cues and a hyporeactivity to visual cues should less easily recognize expressions of emotions than individuals with hyporeactivity to proprioceptive cues and an overreliance on visual cues.

The rest of this Chapter is structured as follows: Section 3.2 describes the collection of the EMBODI-EMO database. Section 3.3 describes its validation by a TD population. Section 3.4 analyzes the relationships between emotion recognition scores and the profiles of our 19 participants with ND. Section 3.5 presents an overall conclusion for the chapter.

3.2 The EMBODI-EMO Database of Facial and Bodily Expressions of Emotions with Different Embodiments

3.2.1 Design

We want to evaluate the effect of different embodiments and different combination of channels, facial expressions and bodily expressions, in the recognition of emotions. The available databases did not appropriately satisfied some conditions: (1) controlling different combination of bodily and face expressions, and (2) their emulation through different embodiment platforms (see 3.1), so we decided to create our own database. We recorded different types of expressions with respect to the channels of expression: (1) monochannel: the emotion is expressed only through one communication channel (e.g., on the face or on the body), and (2) multichannel: body and face express the same emotion in a synchronous manner.

In our work, we used four basic emotions: anger, happiness, fear, and sadness. We selected these four emotions for several reasons. In particular, these emotions are part of the six basic emotions proposed by [Ekman and Friesen, 1984] and are widely documented in the literature. This permits us to have reliable references to design their expressions. Moreover, we can also compare our work with previous studies on how these emotions and their expressions are recognized by TD and ASD individuals. Furthermore, these emotions have been chosen because of their interest and accessibility to individuals with ND. Indeed, while designing the database, we asked caretakers in charge of individuals with ND which emotions would be relevant for training, and they indicated that emotions such as “disgust" and “contempt" were too complex for our participants.

We used four different embodiments to create our emotional video database. They can display different levels of visual information reliability (low vs high completeness) to specify each emotion expressivity, as shown in Figure 3.1:

- Nao (Softbank Robotics): a mini-humanoid robot with a static face (see Appendix B.1).
- Zeno (R25, Robokind): a mini-humanoid robot with a silicon-made actuated skin face,

which enables it to express emotions on its face (see Appendix B.2).

- Mary (MARC, LIMSI-CNRS): a female humanoid virtual agent (see Appendix B.3).
- Human: a female human.

Videos begin with a neutral posture and end with the apex of the expression of emotion. All embodiments occupy the same visual space in the video frame (feet and top of the head were at the same place for each embodiments). The physical platforms (Nao, Zeno, and Human) were placed in front of a white wall, and the virtual agent (Mary) was displayed in a white background. All of the videos are one to two seconds long.

All embodiments perform the same movements. We adapted the animations in order to have similar movements between different embodiments. This was also done in similar studies about bodily expression of emotion [Coulson, 2004, Erden, 2013]. However, several difficulties in movement adaptation were encountered, especially due to different numbers of joints, dynamics of animation, and different balances of the embodiments. We nevertheless did our best to adapt movements through different embodiments to achieve homogeneity.

3.2.2 Facial animation

The embodiments (except Nao) have facial features, which enable them to display prototypical expressions of the four selected emotions (happiness, anger, fear, and sadness). One animation for each emotion was designed based on a single combination of action units (AU) [Ekman and Friesen, 1984] as a function of the capabilities of the embodiments (see Table 3.1). One video was recorded for each emotion for Zeno, Mary, and Human. In [Ekman and Friesen, 1984], the authors propose prototypical expressions of the basic emotions, even if they insist on the fact that there are numerous expression for each emotion. For example, there would be more than 60 ways to express anger with a facial expression. We only recorded one expression for each emotion because (1) we based our animation on the single combination of AU described by [Ekman and Friesen, 1984]; and (2) we were limited by the ability of Zeno to perform facial expressions of emotions.

Table 3.1: Emotion representation by action units (AU), and the ability of Zeno to perform them. Mary (virtual agent) and Human were able to perform all of these AU.

Emotion	AU combination	Zeno
Anger	AU4 - Brow Lowerer	yes
	AU7 - Lid Tightener	yes
	AU10 - Upper Lip Raiser	no
	AU17 - Chin Raiser	no
	AU23 - Lip Tightener	no
	AU24 - Lip Pressor	yes
Happiness	AU6 - Cheek Raiser	no
	AU12 - Lip Corner Puller	yes
	AU25 - Lips Part	yes
Fear	AU1 - Inner Brow Raiser	yes
	AU2 - Outer Brow Raiser	yes
	AU4 - Brow Lowerer	no ∗
	AU5 - Upper Lid Raiser	yes
	AU20 - Lip Stretcher	no
	AU25 - Lips Part	yes
	AU26 - Jaw Drop	yes
Sadness	AU1 - Inner Brow Raiser	yes
	AU4 - Brow Lowerer	no ∗
	AU15 - Lip Corner Depressor	yes

∗ Zeno was only able to have its brow up or down and could not display AU1 and AU4 at the same time.

3.2.3 Body animation

Whereas prototypical facial expressions of basic emotions are documented in the literature, this is not much the case for postural expressions of emotions, especially when it comes to dynamic expressions, various embodiment and their combination with facial expressions. We thus decided to test several bodily expressions for each emotion so that we had more chance to end up with at least one bodily expression being well recognized across the different embodiment that we study. To design our body animations, we inspired from the Bodily Expressive Action Stimulus Test (BEAST) database [De Gelder and Van den Stock, 2011]. This database contains 254 static pictures of whole body expressions from 46 actors expressing four emotions (anger, fear, happiness, and sadness). We selected three images for each emotion to create our animations (anger: the BEAST images, *F03AN, F11AN, F22AN* inspired the design of the animations

AN1, AN2, AN3, respectively; fear: *F04FE, F16FE, F32FE* were chosen for *FE1, FE2, FE3*, respectively; happiness: *F02HA, F06FA, F22HA2* were chosen for *HA1, HA2, HA3*, respectively; sadness: *F02SA, F04SA, F24SA2* were chosen for *SA1, SA2, SA3*, respectively). To animate movement of the whole body from a neutral posture to the posture depicted by BEAST images, we performed a linear interpolation. Although it is recognized that a linear interpolation does not represent human movement [Park and Schowengerdt, 1982], it nevertheless was the easiest way to create our animations. We did not have the possibility to record the movements using motion capture and the processes to transfer recorded biological movements on the three platforms Nao, Zeno and Mary would have been complex to achieve. The velocity of the movement was based on observations of bodily expression in humans [Wallbott, 1998], a greater velocity in movement for anger and fear, and low dynamics for happiness and sadness. Three animations were designed for each emotion on Nao, Zeno, Mary, and Human.

3.2.4 Body and facial animation

We combined each facial animation with the three corresponding body expressions of the same emotion, resulting in three animations for each emotion for Zeno, Mary, and Human.

3.2.5 Content of our database

Our database contains 96 videos:

- 12 videos of facial expressions (four emotions $\times$ three embodiments),
- 48 videos of body expressions (three variations $\times$ four emotions $\times$ four embodiments),
- 36 videos of body and facial expressions (three variations $\times$ four emotions $\times$ three embodiments).

3.3 Validation of the EMBODI-EMO Database with TD individuals

3.3.1 Protocol

We evaluated our database with a TD population. To evaluate the videos, we set-up an online questionnaire. For each video, participants were asked to rate the emotion they recognized (i.e., anger, happiness, fear, or sadness) using a five-point Likert scale (1 = totally disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, and 5 = totally agree). Participants had to evaluate the emotions in different conditions, defined by the combination of channels used in the videos: the facial-only (FO) condition; the body-only (BO) condition, and the body and facial (BF) condition. First, participants evaluated the FO condition, then the BO condition, and finally, the BF condition. Within conditions, the videos were in a randomized order and combined all embodiments. We selected this order because evaluating the BF condition before the FO and BO conditions could have induced a learning effect. In addition, due the characteristics of the system used for the questionnaire, we were not able to randomize the videos for each participant. Therefore, each participant had the same sequence of videos to evaluate. Participants were recruited through mailing lists from our universities (ENSTA-ParisTech, LIMSI-CNRS, and Université Paris-Sud - UFR STAPS). It was an internet form, and no compensation was given.

3.3.2 Participants

Sixty-four participants answered our online questionnaire (31 females, age 28.23 ± 8.31, 62.5% with a technological background, 1 from a non-Western culture).

3.3.3 Data Analysis

First, we observed if the emotions depicted in the videos were correctly recognized by our TD participants. We used the computed the Likert-scale score (LSS) which is the mean value of the scores of the Likert-scales of each video on a five-point scale. Each video had four LSS corresponding to the four emotions. We considered that an emotion was recognized in a video when it obtained a LSS over three. An emotion was correctly identified when the correct emotion was identified in the video. An emotion was confused with another emotion when the emotion identified in the video was not the emotion expressed in it.

Second, we observed the impact of the conditions, the embodiments and the emotions on the recognition of emotions. We computed for each participant the recognition score for each combination of condition, platform and emotion. We first performed a repeated-measure ANOVA on participants' recognition score with three factors: *conditions* with 3 categories (FO, BO and BF), *embodiments* (Zeno, Mary, Human) and *emotions* (anger, fear, happiness, sadness). As Nao was not evaluated in the FO and BF conditions, we excluded it from the evaluation. To observe its effects, we performed a second repeated-measure ANOVA on participants' score only in the BO condition with two factors: *embodiments* (Nao, Zeno, Mary, Human) and *emotions* (anger, fear, happiness, sadness). Post-hoc analyses were performed with Tuckey's HSD test. We used Statistica version 13 to perform the analyses.

3.3.4 Validation of the Videos

FO Condition

In the FO condition, all of the emotions are correctly identified for the embodiments, see Figure 3.2. Fear expressed only through facial features was confused with sadness on Mary. Regardless of this fact, the LSS for fear ($M = 3.42$, $SD = 1.36$) was greater than the LSS for sadness ($M = 3.20$, $SD = 1.48$) on Mary.

BO Condition

The LSS of the videos in the BO condition can be seen in Figure 3.3. All the emotions were recognized on Nao, and no confusion was present, contrary to the other embodiments. On Zeno, only sadness was recognized and not confused with the other emotions. For example, the AN3 (anger) behavior was confused with the happiness emotion. On Mary and Human anger, fear,

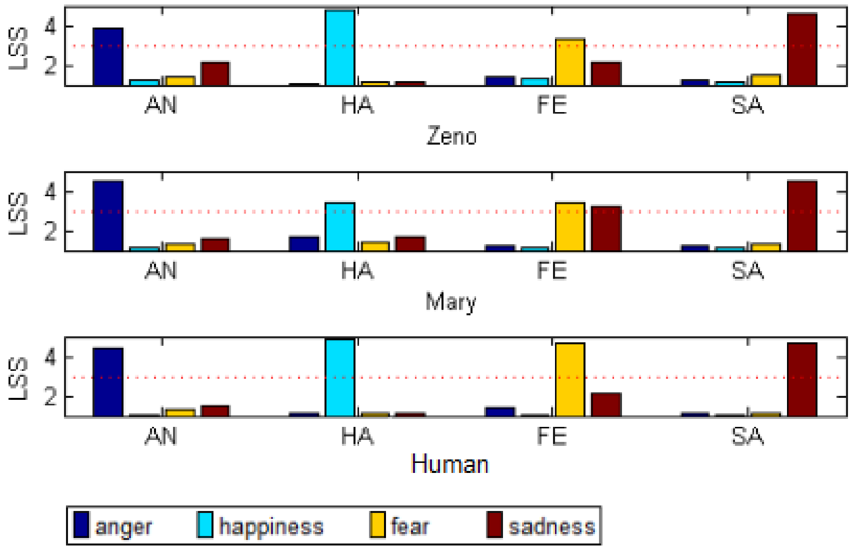


Figure 3.2: Emotion recognition and confusion Likert-scale score (LSS) for each facial animation (facial-only (FO) condition). For each animation (AN# (anger); HA# (happiness); FE# (fear); SA# (sadness)), the mean LSS for each emotion (anger, happiness, fear, sadness) is displayed.

and sadness emotions were correctly identified. However, on Mary, happiness was confused with anger for HA1 (happiness), and not identified for HA2 (happiness) and HA3 (happiness) (the score for anger was greater than the score for happiness for these two animations). On Human, HA2 and HA3 were correctly identified, and HA1 was identified as showing none of the four emotions.

BF Condition

The LSS of the videos in the BF condition are shown in Figure 3.4. Out of the 36 videos presented in the BF condition, 35 videos were correctly identified and no confusion was present. The animation FE3 for Zeno was not well identified. Overall, fear animations of Zeno were weakly recognized (LSS of fear in fear animation was $M = 2.99$, $SD = 1.49$ on Zeno), and it was confused with anger (LSS of anger in fear animation was $M = 2.60$, $SD = 1.53$ on Zeno).

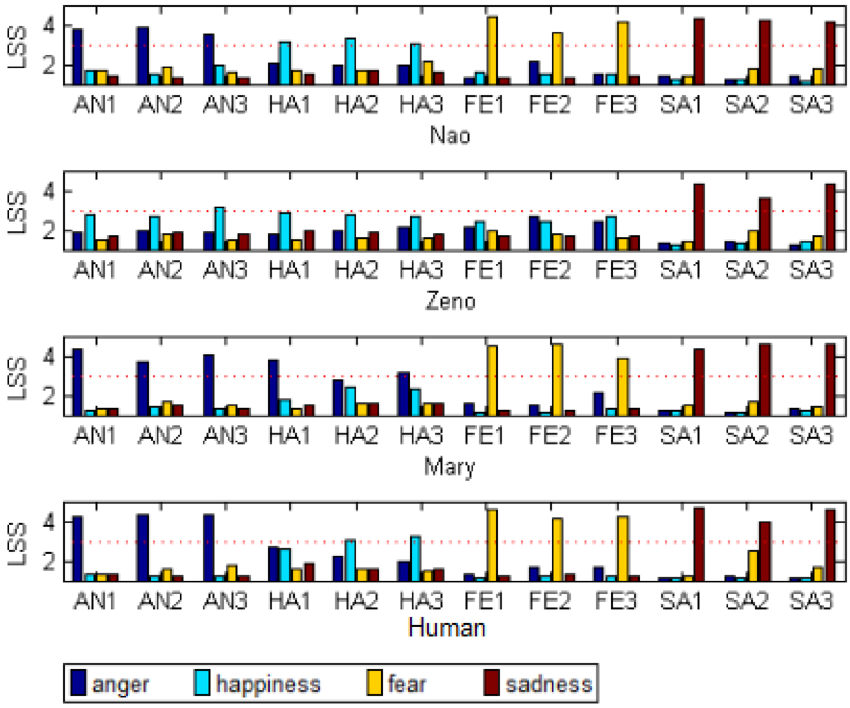


Figure 3.3: Emotion recognition and confusion Likert-scale score (LSS) for each body animation (body-only (BO) condition). For each animation (AN# (anger); HA# (happiness); FE# (fear); SA# (sadness)), the mean LSS for each emotion (anger, happiness, fear, sadness) is displayed.

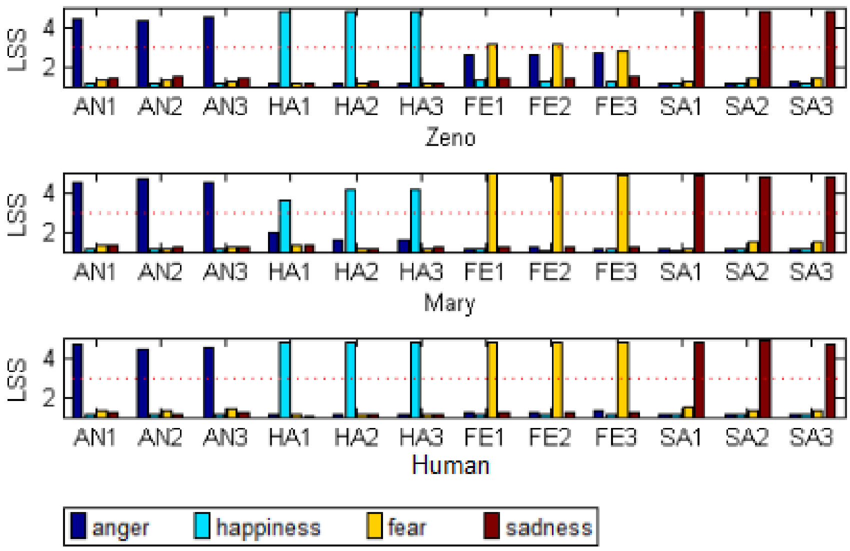


Figure 3.4: Emotion recognition and confusion Likert-scale score (LSS) for each congruent combination of body and facial animation of the same emotion (AN# (anger); HA# (happiness); FE# (fear); SA# (sadness)), the mean LSS for each emotion (anger, happiness, fear, sadness) is displayed.

3.3.5 Impact of the Conditions, the Embodiments (Nao Ecluded) and the Emotions on the Recognition of Emotions

Main Effects of the Conditions

Results showed significant main effects of the conditions ($F(2, 126) = 191.52$; $p < 0.001$) see Figure 3.5a. Emotions in the BF condition were the most recognized ($M = 4.64$, $SD = 0.84$), followed by emotions in FO condition ($M = 4.26$, $SD = 1.06$); emotions in the BO condition were the least recognized ($M = 3.45$, $SD = 1.55$). These results indicate that the combination of both body and facial features enhance the recognition of an emotion and that facial features play a more important role in emotion recognition of emotion categories than body features.

Main effects of the Embodiments

Results showed significant main effects of the embodiments ($F(2, 126) = 228.65$; $p < 0.001$), see Figure 3.5b. The emotions were better recognized on Human ($M = 4.60$, $SD = 0.81$) than on Mary ($M = 4.18$, $SD = 1.15$) and Zeno ($M = 3.82$, $SD = 1.45$). Zeno had the smallest recognition score. These results show the impact of the embodiments on emotion recognition.

Main effects of the Emotions

Results showed significant main effects of the emotions ($F(3, 189) = 72.17$; $p < 0.001$), see Figure 3.5c. Sadness was the best recognized emotion ($M = 4.70$, $SD = 0.60$), followed by anger ($M = 4.18$, $SD = 1.19$). Fear ($M = 3.90$, $SD = 1.42$) and happiness ($M = 4.00$, $SD = 1.34$) were the least recognized emotions.

Interaction effects between the Conditions and the Embodiments

Results showed significant interaction effects between the conditions and the embodiments ($F(4, 252) = 76.82$; $p < 0.001$). Post-hoc analyses on conditions and embodiments interaction (see Figure 3.6) suggest that recognition scores on Human are statistically different between the FO and BO conditions and also between the BO and BF conditions (respectively $p < 0.001$ and $p < 0.001$), but not between the FO and BF conditions ($p = ns$). Identically, recognition

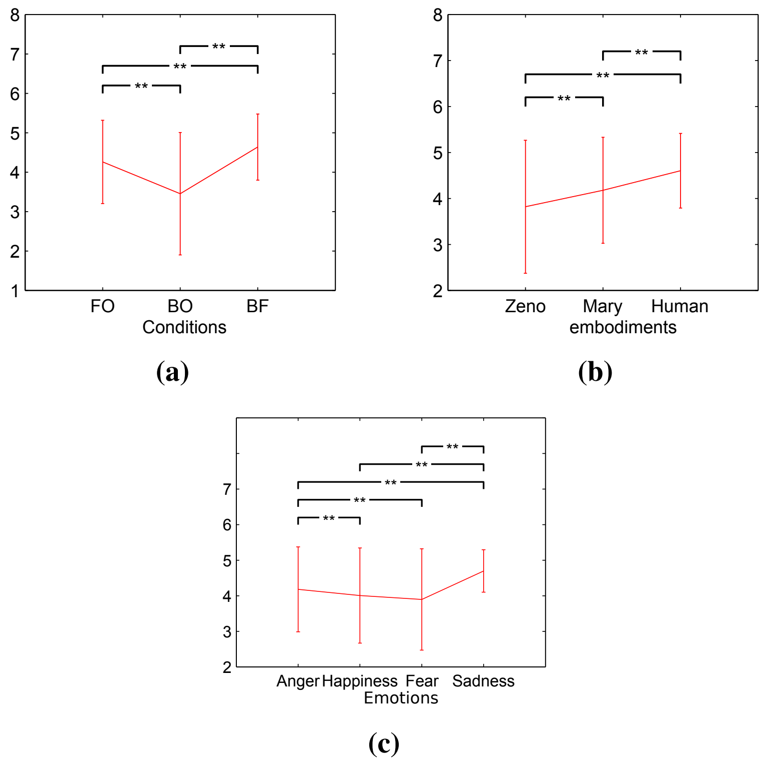


Figure 3.5: Correct recognition scores for the (a) conditions, (b) embodiments and (c) emotions. ** $p < 0.001$.

scores on Zeno are statistically different between the FO and BO conditions and also between the BO and BF conditions (respectively $p < 0.001$ and $p < 0.001$), but not between the FO and BF conditions ($p = ns$). On Mary, all conditions were statistically different in-between them, and the recognition score was the highest in the BF condition ($M = 4.73$, $SD = 0.61$). This indicates that when displayed on Zeno and Human embodiments, the body features did not bring more information to participants to help them to recognize the emotions. On the contrary, when displayed on Mary, body features in addition to facial features helped to recognize the emotion.

Interaction effects between the Conditions and the Emotions

Results showed significant interaction effects between the conditions and the emotions ($F(6, 378) = 46.06$; $p < 0.001$). Post-hoc analyses (see Figure 3.7) informed us that for each emotion, the recognition scores are significantly different between the FO and BO conditions, and between the BO and BF conditions. Between the FO and BF conditions, only fear

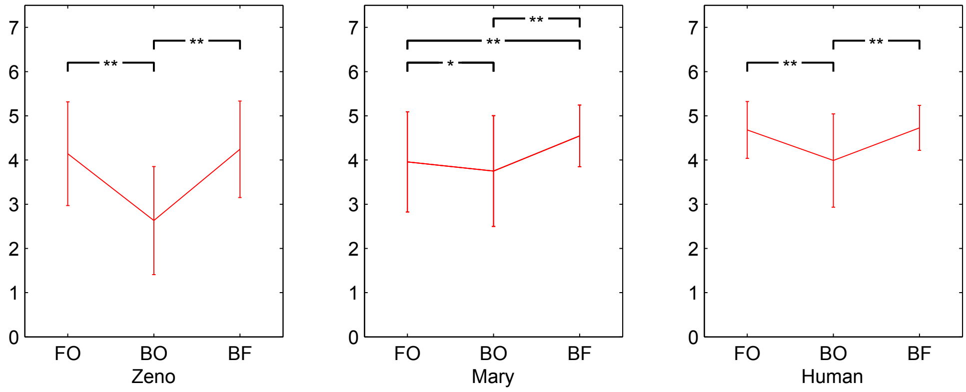


Figure 3.6: Correct recognition scores of each embodiment in each condition. * $p < 0.01$ ** $p < 0.001$.

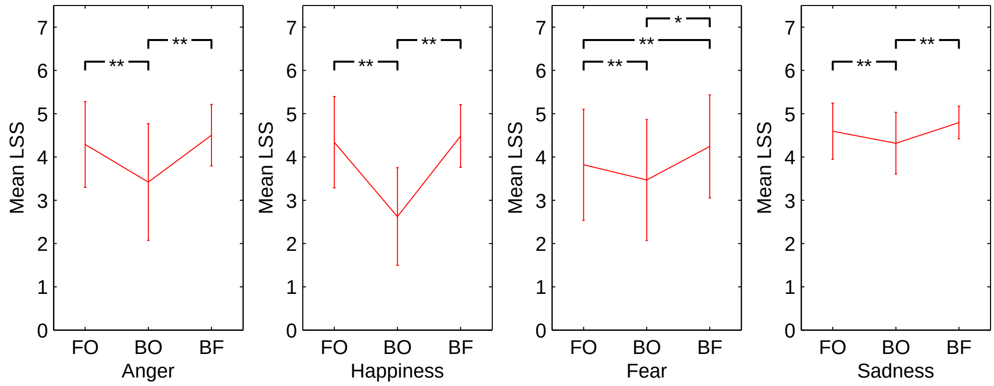


Figure 3.7: Correct recognition scores of each emotion in each condition. * $p < 0.01$ ** $p < 0.001$.

had significantly different scores ($p < 0.001$). We can also observe on Figure 3.7 that for happiness, the score decrease was the strongest in the BO condition. This suggests that fear was the only emotion in which body features improved the recognition level from the FO condition to the BF condition. All the emotion were less recognized in the BO condition than in the FO and BF conditions. For happiness we can observe that the recognition score decrease was more important than the other emotions. This difference might be explained by the fact that in the BEAST Database [De Gelder and Van den Stock, 2011] (i.e., the database we used to inform the design of the animations of the body), happiness was the least recognized emotion.

3.3.6 Impact of the Embodiments (Nao Included) and the Emotions on the Recognition of Emotions in Condition BO

Main effects of the Embodiments

Results showed significant main effects of the embodiments ($F(3, 189) = 251.10$; $p < 0.001$). Post-hoc analyses indicated significant differences between recognition scores across all embodiments, excepted between Nao and Mary ($p = ns$), indicating that Mary and Nao showed similar recognition scores, even if they have a different embodiment (ie., virtual vs. real embodiment, no facial features for Nao).

3.3.7 Conclusion

Out of the 96 videos, emotions were correctly identified in 82 videos (12 out of 12 in the FO condition, 35 out of 48 in the BO condition, and 35 out of 36 in the BF condition). Emotion recognition results were higher in conditions in which the face expressed emotion (FO and BF conditions) than in the condition where only the body expressed the emotion (BO condition). This highlights the importance of facial expressions for recognizing the category of emotion, as observed by [Buisine et al., 2014] and [Meeren et al., 2005]. In addition, body features have been shown to highlight and intensify facial features in emotion recognition in TD individuals [Clavel et al., 2009, De Gelder et al., 2015]. Sadness was the easiest emotion to recognize, followed by anger. Happiness and fear were the most difficult to recognize. In the BO condition, the LSS of each emotion followed the same pattern found in the BEAST Database [De Gelder and Van den Stock, 2011] (i.e., the database we used to develop the body animations). Sadness was the best recognized emotion, followed by fear, anger and happiness. In [Russell, 1994], the authors reviewed eight studies on the recognition of the six basic emotion in Western and non-Western cultures, and provided their mean rate of emotion recognition. As there was only one out of 64 of our participant in our study that comes from an non-Western culture, we compared our results with the Western culture rate. In [Russell, 1994], they found that, while expressed by the face, happiness was the most recognized emotion, followed by sadness, anger and fear. In our study, in the FO condition, we found that sadness ($M = 4.59$; $SD = 0.65$) was the best recognized, followed by happiness($M = 4.33$; $SD = 1.05$), anger($M = 4.26$; $SD = 0.99$) and fear($M = 3.8$; $SD = 1.28$).

3.4 Emotion Recognition by Individuals with ND

In this section, we investigate the relationship between the individual sensory preferences defined in Chapter 2 for our participants with ND and their abilities to recognize emotions expressed in the videos of our EMBODI-EMO database (validation of the **H1'** hypothesis). We also consider the relationship between the age of our participants with ND and their abilities to recognize emotions.

3.4.1 Protocol

The evaluation of emotion recognition by individuals with ND was done on a subset of the videos of the database. We agreed with the caretakers that evaluating 96 videos would be a difficult task, especially for individuals with ND. We selected the videos with the highest recognition scores obtained with TD individuals for each emotion in each condition for each embodiment, resulting in a 40 videos subset. The selected videos were as follows:

- FO condition: all of the videos were selected
- BO condition:
  - Nao: AN3, FE3, HA2, SA1
  - Zeno: AN2, FE1, HA1, SA1
  - Mary: AN2, FE3, HA3, SA1
  - Human: AN1, FE2, HA3, SA1
- BF condition:
  - Zeno: AN3, FE1, HA2, SA1

- Mary: AN2, FE3, HA2, SA1
- Human: AN2, FE3, HA2, SA2

3.4.2 Method

We developed a graphical interface to evaluate individuals' skills to recognize bodily and facial expressions of emotion on different embodiements. We designed a tactile computer game, as shown in Figure 3.8. A video was shown to the participants. Then, the participants were asked to categorize the emotion recognized in the video with a forced choice. The choices were displayed at the bottom of the screen, with Picture Exchange Communication System (PECS) images [Bondy and Frost, 1994] and labels for anger, happiness, fear, and sadness. The participant could replay the video. The possibility to skip to the next trial without answering the question was not allowed because we feared that some of the participants would have skipped all of the videos without trying to recognize any emotion at all. However, if we noticed that the participant was lost or verbalized his/her inability to answer, the experimenter could skip the current video and move on to the next video. The software recorded the time between the launch of a video and the selection of an emotion category. We divided our 40 videos into four sessions of ten videos to be able to take breaks between these sessions. The task was quite demanding and required high cognitive levels of concentration from the participants, hence, these breaks were necessary. A counter was displayed at the top of the screen to inform the participant of his/her progress within the current subset of videos. Before performing the game, the procedure was explained to each participant to ensure that the instructions were correctly understood. During the tests, the experimenter and the caretaker were as neutral as possible so as not to interfere with participants' recognition skills. They did not help the participants or correct them in the choice of their answers. The order of the presentation of the conditions and of the videos followed the same principle that we used for the TD participants. First, participants evaluated the FO condition, followed by the BO condition, and finally, the BF condition. Within conditions, the videos were presented in a randomized order and combined all embodiments, i.e., the Nao, Zeno, Mary, and Human, animations were randomized. The order of the videos

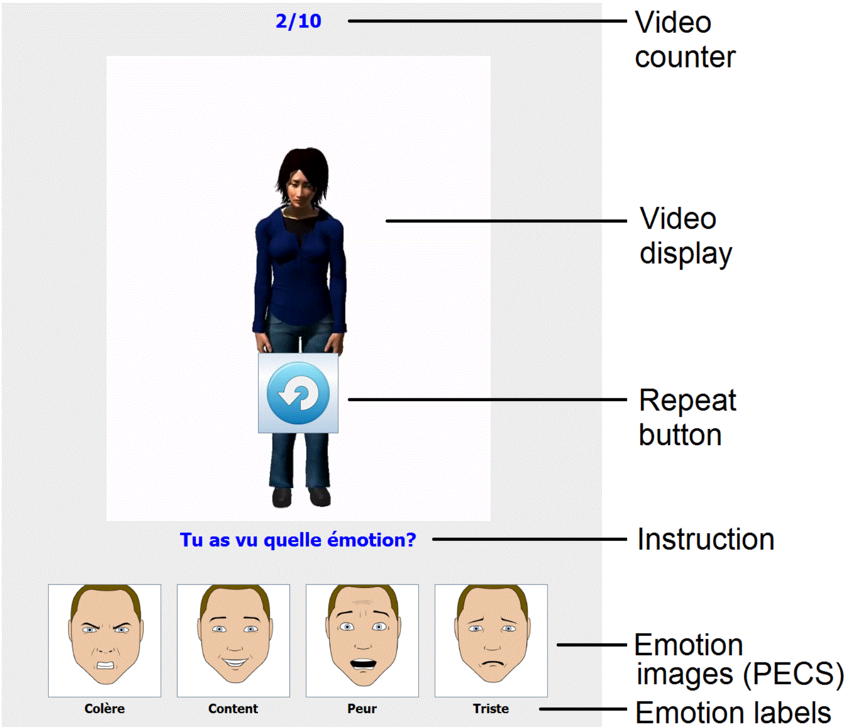


Figure 3.8: Screenshot of the game at the end of a video. Emotion labels (written in French) are, from left to right: anger, happiness, fear, and sadness. The instruction, Tu as vu quelle émotion ?" means “Which emotion did you see?"

was randomized for each participant.

3.4.3 Participants

All participants from our subject pool (Section 1.3.2) performed the experiment, except for participants AD3 and CH11: they were not present in the care facilities when the experiment was conducted.

3.4.4 Data Analysis

We observed the impact of the conditions, the embodiments and the emotions on the recognition of the emotions. For each video, the participants got a correct or a wrong answer. We observed if the groups (G1, G2 and G3) and the age (children and adults) of the participants had an impact on emotion recognition. First, we performed a repeated-measure ANOVA on participants' answers with three factors considered for the analysis: *conditions* with 3 categories (FO, BO and BF), *embodiments* (Zeno, Mary, Human) and *emotions* (anger, fear, happiness, sadness). As Nao was not evaluated in the FO and BF conditions, we excluded it from the evaluation. To observe its effects, we performed a second repeated-measure ANOVA on participants' answers only in the BO condition with two factors considered for the analysis: *embodiments* (Nao, Zeno, Mary, Human) and *emotions* (anger, fear, happiness, sadness). Post-hoc analyses were performed with Tuckey's HSD test. The significance threshold was set to $p < 0.05$. We used Statistica version 13 to perform the analyses.

3.4.5 Results

Participants' Score and Behavior during the Experiment

Below, we describe the behavior and scores of each participant with ND. Participants AD3 and CH11 were not in the care facilities during the sessions. A brief description of the participants was provided previously in Section 1.3.2. Most of the participants found that the task was difficult.

AD1 (G2 group): In the BO and BF conditions, she chose the answer “Happy" most of the time. Nevertheless, she seemed to understand the game's instructions.

AD2 (G1 group): She selected happiness for 38 of the 40 videos. When we noticed this trend during the break after the first session, we checked to see if she understood the task and asked her to describe to us the movements she saw on the screen. She was able to describe them and understood the task. Nevertheless, she always chose the happy emotion, except for the last two videos where she chose the correct emotions, which were not happiness. For the duration of the experiment, she showed great enthusiasm, but started to be more quiet at the end, which is when she chose the correct emotions. After careful discussion with the caretakers, we determined that she was really happy to perform the task with us, as she liked to be the center of attention. Moreover, she always smiled or said statements with a happy connotation. Her answers may thus have been influenced by her own affective state (happy).

AD4 (G1 group): She showed great enthusiasm for performing the task. She had difficulties with the FO condition (16.7% good recognition), but displayed higher scores in the BO and BF conditions (BO: 56.3%; BF: 91.7%). She obtained the best scores among the G1 group.

AD5 (G3 group): He was recovering from a flu, but performed the task with care. However, it was clear that he was really tired at the end of the task. He had great difficulties in recognizing Zeno's emotions (0% of good recognition).

AD6 (G1 group): He had difficulties understanding the task. The caretaker had to help him touch the screen. Since he suffers from echolalia, we think that he only repeated what he was asked when he did not know what to choose (i.e., when he did not seem to answer, he was asked, "What do you think the character shows? Anger? Happiness? Fear? Sadness?" and he repeated, "Sadness," as it was the last word of our sentence). However, he answered some of the items without hesitation, showing that he nevertheless partially understood the instructions.

AD7 (G3 group): He was always struggling while choosing an emotion to evaluate the videos (a common symptom in ND). It took him a long time to perform the task.

CH1 (G2 group): He was really reluctant to perform the task. This can be explained by the facts that (1) he was coming down with a flu, (2) a recent change in his referent caretakers occurred, which made him reluctant to many tasks.

CH2 (G3 group): He understood the task and performed it calmly. Overall, he had a good recognition score (67.5%). He had more difficulties with the BO condition than with the other conditions (31.3%) and had more difficulty with Nao than with the other embodiments (25%).

CH3 (G1 group): He understood the task. However, at the end, he seemed to "accelerate" the procedure and not properly choose the correct image. He showed good recognition scores in the FO condition and with the real human embodiment.

CH4 (G3 group): He succeeded in correctly recognizing the emotions from 39 of the videos. However, his process in understanding the emotion in the BO condition is noteworthy: after looking at the videos, he did not directly recognize the emotions. He imitated the movement in front of the computer saying, “How do I feel if I do that" and succeeded to understanding them. He presented the best cognition capacity among all of the participants.

CH5 (G1 group): We were unsure of her capacities to perform the task, as she is non-verbal and unfamiliar to this kind of task. However, she understood it and showed an interest in the task (she replayed most of the videos). She had a good emotion recognition rate in the FO condition and for the Nao and Zeno embodiments.

CH6 (G3 group): He easily understood the game. This can be explained by the fact that a similar game is performed as part of his weekly therapy. However, the task put him in a state of frustration when he was not able to provide the correct answer, especially in the BO condition.

CH7 (G2 group): He understood the game but nevertheless expressed some difficulties to recognize the emotions.

CH8 (G1 group): He had difficulties to understand the task and to take decisions.

CH9 (G2 group): She understood the task. However, she had some troubles to use the touchscreen.

CH10 (G1 group): He understood the game and liked to repeat the videos. He sometimes imitated the animation in front of the screen.

CH12 (G2 group): He understood the game, and recognized the emotions. But for some animations, he selected apparently randomly the answer, without any known reasons.

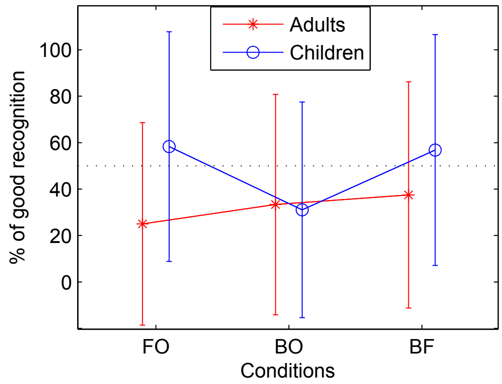


Figure 3.9: Recognition scores for adults and children with ND in each condition

Impact of the Conditions, the Embodiments and the Emotions on the Recognition of Emotions in children with ND

At first, no main effects of the groups of the participants were found. We observed a weak difference between children and adults recognition scores ($F(1, 10) = 4.2$; $0.05 < p < 0.1$): children recognized the emotions better (48.74%) than the adults (31.94%). By looking more closely to the scores of our participants (see Figure 3.9), and with the observation of the behaviors of the participants during the recognition task, the adult participants seemed to have more difficulties to answer than the children. The recognition scores of the adults were below 40% in all conditions, differently to the children. Children had a recognition level of 58.33%, 31.06%, and 56.81% in the FO, BO and BF conditions respectively. Adults had a recognition level of 25%, 33.33%, 37.5% in the FO, BO, and BF conditions respectively. We hypothesized from these results and from the behaviors displayed during the sessions that adult participants mostly answered randomly to the questionnaire, whereas children did not. We continue the analyses only with the children participants.

Main effect of the Conditions (Nao excluded)

A significant main effect of the condition on the recognition scores was found in children with ND ($F(2, 16) = 20.53$; $p = 0.001$). Post-hoc analysis showed us that scores were statistically different between the FO and BO conditions ($p < 0.001$), and between the BF and BO conditions ($p < 0.001$). Emotions in the BO condition ($M = 31.06\%$; $SD = 46.45\%$)

were less recognized than in the FO ($M = 58.33\%$; $SD = 49.49\%$) and BF ($M = 56.81\%$; $SD = 49.72\%$) conditions. This indicates that the emotions only expressed with body features were more difficult to recognize for children with ND.

Main effect of the Embodiments (Nao excluded)

No significant main effect of the embodiments were found in children with ND.

Main effect of the Emotions (Nao excluded)

No significant main effect of the emotions were found in children with ND.

Main effect of the Groups (Nao excluded)

We observed a weak effect of the groups ($F(2, 7) = 3.45$; $0.05 < p < 0.1$) on emotion recognition. Children from group G1 had a score of $M = 29.86\%$; $SD = 45.92\%$, children from group G2 had a score of $M = 53.33\%$; $SD = 50.03\%$ and children from group G3 a score of $M = 75.00\%$; $SD = 43.61\%$. Children from group G1 had the lowest score and children from group G3 the highest as we expected.

Interaction effect of the Groups and the Conditions (Nao excluded)

We found a significant interaction effect of the conditions and groups ($F(4, 14) = 3.76$; $p < 0.05$), see Figure 3.10. A post-hoc analysis showed that for children of group G3, scores in the BO condition is statistically different of scores in the FO ($p < 0.05$) and BF ($p < 0.05$) conditions. Children from group G3 showed more difficulties to recognized emotions in the BO conditions than in other conditions (which all use facial features). We observe this pattern in children from groups G1 and G2 (see Figure 3.11), but it is not significant. The addition of facial features enabled children from group G3 to recognize significantly more the emotions. In the BF condition, participants from group G3 had 100% scores. Children relying more on visual features and less on proprioceptive features seemed to extract more the information of the face, differently to participants relying less on visual features and more on proprioceptive features.

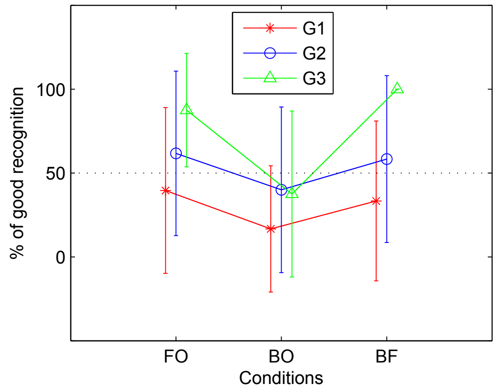


Figure 3.10: Recognition scores of the three groups (children participants only) in all conditions

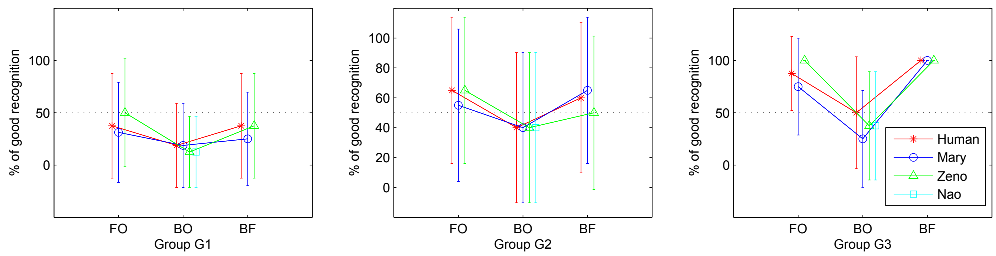


Figure 3.11: Recognition scores for each embodiment in each condition of the three groups (children participants only)

Results in Condition BO (Nao included)

No significant results were found in the BO condition when Nao was included.

3.4.6 Conclusion

The results of the task seemed to be led by the motivation and/or condition of the participant. The task was described as difficult by the participants, showing their involvement and efforts in performing the task. There seems to be a relationship between the difficulty reported and the inability to properly recognize emotions. This may also explain the overall low emotion recognition scores. In [Rump et al., 2009], the authors tested the performance in emotion recognition of children, teenagers and adults with ND and of a control group. They found that the performance level was similar in all age group in individuals with ND, in contrast to TD individuals, for who the performance level is the best in adults. In our study, children performed better than the adults. This could be explained by the fact that they often perform similar tasks with their caretakers (which is not the case for the adult participants we worked with). Adults showed recognition scores close to random. By removing the adults' results from the analysis, we were able to observe that children from group G1 had the lowest score and the children of group G3 the highest score as we expected in our hypothesis **H1'**. We also observed that children with an overreliance on visual cues and a hyporeactivity to proprioceptive cues seemed to extract more information from facial features than the other participants. This goes in the direction of the findings of [Atkinson, 2009]: adults with ASD that were less efficient in perceiving coherent motion were less efficient in recognizing emotions. In [Uljarevic and Hamilton, 2013], the authors proposed a meta-analysis regrouping 48 papers published between 1989 and 2011 testing emotion recognition in individuals with ASD. They observed that happiness was the most recognized emotion and fear the least recognized. In our results, no effects of the embodiment or the emotions were found, differently to TD individuals. Similarly to TD individuals, conditions including the facial features were better recognized than when only body features were used in the animation of the emotions. Emotions from only body channel were difficult to recognize by children with ND.

3.5 General Conclusion

In this Chapter, we aimed to establish a way to evaluate the capacity of our participants to recognize bodily and/or facial (in isolation or combined) expressions of emotions with respect to their capacity of integration of visual and proprioceptive cues.

First, we developed a database (EMBODI-EMO) of 96 videos expressing four emotions (anger, happiness, fear, and sadness) on four embodiments (human, virtual, robots) with different types of channels (with or without facial expressions, combined or not with body movements). Second, we evaluated the EMBODI-EMO database with TD participants. For the 96 videos, the emotions were correctly identified by the TD participants in 82 videos. As in [Meeren et al., 2005, Buisine et al., 2014], we found that the combination of body and facial expressions of emotions enhanced emotion recognition for TD participants. A smaller subset of the database, composed of 40 selected videos with the best recognition scores for each emotion, for each platform, and for each condition, was selected to evaluate individuals with ND. Indeed, the caretakers stated that evaluating 96 videos would have been a difficult task for individuals with ND. Finally, we evaluated the subset of 40 videos of the EMBODI-EMO database with our participants with ND. We observed that most of our participants found the task difficult. Adults showed worse scores than children. We posit that this difference comes from the different tasks our participants perform in their care facilities. Children do similar tasks as the matching game and work more on emotions than the adults we work with. As the adults answered poorly, we removed them from the analysis. By doing the analysis on only the children participants, we observed that children with an overreliance on visual cues and a hyporeactivity to proprioceptive cues performed better than children with an overreliance on proprioceptive cues and a hyporeactivity to visual cues, which validates our hypothesis **H1'**.

These results provided evidence that integration of visual and proprioceptive cues of an individual with ND influence his/her abilities in recognizing emotions. Participants with ND relying on proprioceptive information had lower recognition scores than the other participants with ND. In future work, these findings can help us to develop personalized interaction. Designers of Human-Machine Interaction for users with ND may thus select the channels (e.g., facial expression) or the type of embodiment (e.g. robots vs. virtual agent vs. human) depending on the user's visual and proprioceptive profile.

Participants with ND found the task difficult. As observed in their behaviors, AD2 and CH4 had good observation strategies, but had impairments in understanding emotions, similar to the observations in another study [Celani et al., 1999]. They properly observed the movements displayed in the videos. CH4 succeeded in understanding the emotions using imitation, but AD2 was overwhelmed by her own internal state. We could have chosen to include a context for our videos, as in [Begeer et al., 2006]. Here, the authors observed that children with ASD had more difficulties recognizing emotions when they were out of context, but they recognized emotions better when contextual information was provided. In their study, children were shown four pictures of faces and had to match two of them. In the condition with context, the experimenter prompted the child through different sentences such as *“Which two would be most likely to give you a sweet?"* or *“Imagine all of the men in the pictures are teachers. Which two teachers are most likely to tell you off?"* However, in our setup, the inclusion of context for the emotions would have been more difficult. We did not want to induce cues regarding the answer by adding a context or verifying that each participant understood the background story. Several studies pertaining to the recognition of expressions of emotions do not involve any context at all [Baron-Cohen et al., 2001a].

In our software, we displayed images for the selection of emotions; however, this possibly had an impact on our participants. They may have tried to match them to the specific facial expression images shown in the video, rather than the more general concept of the emotion. During the design process of the system, we looked for the best way for participants to choose one of the emotion categories (happiness, angry, fear, and sadness). We searched for a simplified, easy to understand method for all of our participants. Since only a few of our participants were able to read, we could not use only written labels; hence, we chose images in addition to labels. The use of pictograms is widely used by individuals with ASD (PECS, for example [Bondy and Frost, 1994]) to describe the planning of the day or their emotional state. We also imagined showing several types of pictograms of the face, body, and face and body combined. However, we did not find a validated set of this kind of pictograms. The use of faces to describe emotion is also more common and already used by the families and caretakers of the participants in our study. As such, we decided to use PECS [Bondy and Frost, 1994] images of facial expressions of emotions. A limitation of our work is that these expressions of emotions are quite prototypical. Natural and prototypical or acted emotions lead to different expressions and underlying processes [De Gelder and Van den Stock, 2011, Dael et al., 2012]. However, for the postural expressions of emotions, we inspired from the BEAST validated database [De Gelder and Van den Stock, 2011], which also contains expressions of acted emotions. Furthermore, acted emotions are used for emotion recognition in therapy for individuals with ASD, such as for the *Mind Reading* software [Baron-Cohen et al., 2001a].

We also noticed that the task (i.e., choosing a label for the video) was too complex for some of the participants with ND. However, by observing their behavior during the session, we are confident that the instructions were understood. Some of the participants had difficulties understanding or linking an emotion to the video they were watching. In addition, in the results, we observed that children with ND participated better than adults with ND. This may be caused by the fact that children perform this kind of task more often than adults in their care facility. Prior to the task, we tried to avoid this effect by explaining and showing to participants how to perform the task before doing it. However, in future studies, it might be more effective to train subjects with the same system but without the emotions involved. For example, we could ask a participant to perform a complete session of recognizing actions (such as walking, running, jumping, etc.), to make sure they are familiar with the task, and then proceed to the emotion recognition task. Since we wanted to assess a participant's capacity to recognize emotions, we did not want to overexpose them to the emotion videos in order to avoid any learning effects.

We also noticed that the experimenter and caretaker behaviors had an impact on the participants. Indeed, when the subjects played or worked with their caretaker, they generally receive feedback to correct or be encouraged. In our sessions, we did not exhibit this behavior in an effort to not induce any learning effects. Some of the participants showed frustration during the task, as no feedback was provided.

Another limitation is that TD participants and participants with ND may have learned the emotions while doing the experiment. Indeed, the first presentation of emotions during the FO and BO conditions could have been a preamble for the participants to recognize the emotions for the BF condition. However, a learning effect in a database of this size is difficult to avoid.

Chapter 4

Task #2: Greetings with a Robot

In this chapter, we present a first interaction with the Nao robot and the participants. The objective of this first interaction was to introduce the robot as a social partner and to remove the novelty effect of the robot for further experiment. We also wanted the participants to be more at ease with the robot, as individuals with ND can show stress or reluctance to new things. Furthermore, we observed the social behavior of our participants towards the robot, and observed if there was a link with their profiles as described in Chapter 2. The most visual participants showed more free speech than the other participants, and the most proprioceptive participants were the less talkative. There was no statistical evidence that the most visual participants looked more at the robot during the interaction than the other participants, yet we did observe a tendency.

4.1 Introduction

The use of robots in therapy for individuals with ASD has been a great topic of interest over the past years. Indeed, robots have been found to be great partners for facilitating learning, social interaction, and developmental imitation skills. In addition, robots are attractive mechanical systems for people with ASD [Hart, 2005, Kim et al., 2013]. They are appealing and engaging for individuals with ASD because their nonverbal behaviors are quite simple and predictable. Robots allow a user to experience endless social interactive scenarios in which complexities of social situations, of bodily expression of emotions, can be controlled to facilitate the detection of regularities and predictable events.

In this chapter, we describe an experiment where Nao robot (see Appendix B.1) was presented for the first time to all our participants. Their behaviors, during this first interaction with the robot, based on their profiles are analyzed.

The chapter is structured as follows: Section 4.2 describe the protocol of this experiment. Section 4.3 presents the results. Finally, Section 4.4 presents a conclusion for the Chapter.

4.2 Protocol

4.2.1 Objectives

The purpose of the interaction was to introduce the robot to the children and adults with ND for a short duration (up to 2 minutes). Some individuals with ND are reluctant to unusual events and changes in their daily routine. The robot was smoothly introduced so as to avoid fear of the robot. In addition, the authors in [Meltzoff et al., 2010] observed that children who saw the robot act in a social-communicative way were more likely to follow its gaze than those who did not. Hence, we believe that introducing them smoothly the robot as a social partner by showing them Nao in the context of a short greeting task may help the participants to interact with the robot in further experiments.

We also wanted to test if the behaviors of our participants during this initial task were linked to their proprioceptive and visual profiles as described in Chapter 2. To do so, we analyzed the interaction with Nao robot and annotated our participants' social behavior following the items described in Table 4.1.

4.2.2 Method

The scenario was the following: after being seated in front of Nao, the robot waved to the participant and said "Hello, I am Nao. You and I, we are going to be friends." (Figure 4.1). If

Table 4.1: Description of the tracked social behaviors

Smiles, laughter
Speech to Nao (i.e., said hello/goodbye; gave his/her name by himself/herself; answered if he/she wanted Nao to dance): 1. by his/her initiative 2. after being encouraged by his/her caregiver
Waving back gesture
Gaze of the participant to: 1. the robot 2. his/her caregiver 3. other (somewhere else)

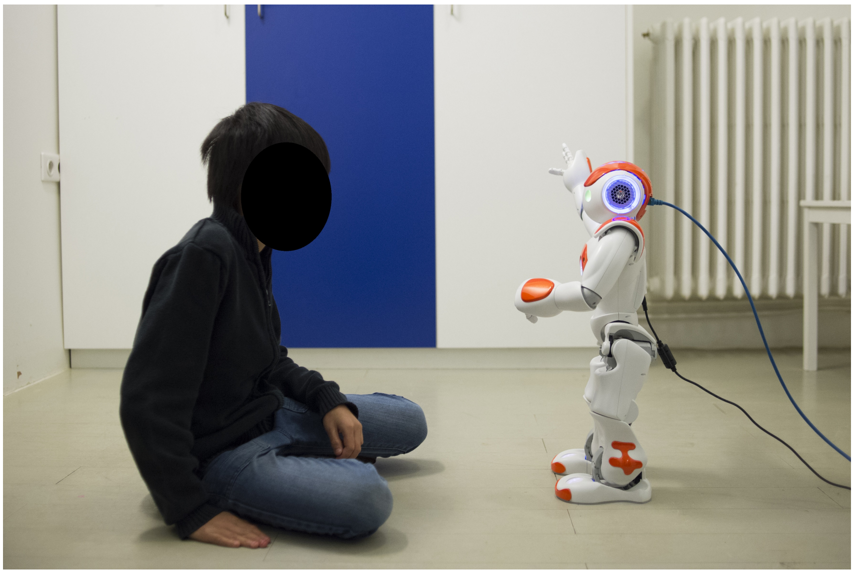


Figure 4.1: Nao greets a child

the participant was verbal, the robot asked for the participant's name. Afterwards, it continued with saying "Hello" followed by the name of the participant, and asked if he/she wanted it to dance for him/her, and then danced (Figure 4.2). During all the experiment, the participant was accompanied by his/her caregiver. The caregiver encouraged the participant to look at and to answer to the robot.

The interaction was simulated using a Wizard of Oz (WOZ) protocol [Kelley, 1983] involving an experimenter in the interaction loop. We used Choregraphe[1], the software provided by Softbank Robotics to program Nao's movements, action and sentences.

[1] http://doc.aldebaran.com/1-14/software/choregraphe/index.html

Figure 4.2: Nao dances for a child

4.2.3 Data Analysis

We analyzed the videos of the interaction with the robot and observed the parameters described in Table 4.1 for each participant. A first coder manually annotated all of the videos of the interaction. A second coder, unaware of the hypotheses of the setup, annotated a 21% random selection of the videos. The Intraclass Correlation Coefficient (ICC) was used to ensure inter-coder reliability. The ICC score was of 0.99, indicating a very good reliability.

AD3 was removed from the statistical analysis as he was becoming more withdrawn from social interaction since a few weeks, and was unwilling to participate in the tasks. CH5 was removed from the statistical analysis of the speech data as this participant is non-verbal. For the gaze and smile behavior analyses, we performed a one-way ANOVA among the groups. We performed Fisher's Exact Test on the speech and gesture behaviors. The significance threshold was set to $p < 0.05$. We used Statistica version 13 to perform the analyses.

4.3 Results

The participants' behaviors are described in Tables 4.2, 4.3, and 4.4. Overall, except for AD6 (G1) and AD3 (G2), participants from all groups looked more than 60% of the time to the robot. No statistical evidence was found about the gazing behavior of our participants and its relation with the groups. However, we can still observe on Figure 4.3 the frequency/percentage distribution of the participants' gaze to the robots among the groups. Participants from group G3 gazed more frequently towards the robot than the two other groups, and participants from group G1 appeared to display fewer gaze to the robot than the two other groups. Participants from all groups smiled during the interaction. No statistical evidence was found on smiling behavior and groups. However, we can notice on Figure 4.3 that participants from group G1 appeared to smile more than participants from groups G2 and G3. The number of participants speaking to Nao by their own initiative was significantly different among groups ($p < 0.05$). Only one participant out of seven from group G1 responded by its own initiative to the robot, when five out of six participants from group G2 and three out of four participants from group G3 spoke to Nao by their own initiative. We can see that participants CH8, CH11, and AD2 from group G1 were the only participants not to respond to the robot with or without encouragement. No statistical evidence was found between waving back gesture and groups. However, we observed that participants from group G1 did not show waving back gesture to the robot, but two participants from group G2 and two participants from group G3 did.

4.4 Conclusion

The presentation of the Nao robot to the ND participants permitted us to introduce it as a social partner. Most of the participants answered to it and some displayed social behavior to it (gaze, smile, speech). This introduction to the robot may help the participants to interact easier with the robot in further experiments, as found in [Meltzoff et al., 2010]. We also removed some of the "surprise" and "novelty" effect of the robot. Some participants showed to be slightly afraid and impressed by the robot. These participants seemed to be reassured at the end of the interaction.

Table 4.2: Participants' percentage of gaze behavior and smiles during the interaction with the robot

G#	ID#	Gaze to Nao	Gaze to Caregiver	Gaze to Other	Smiles
G1	CH3	63.1%	18.17%	18.73%	19.93%
	CH5	81.9%	2.76%	15.34%	0%
	CH8	100%	0%	0%	100%
	CH10	88.52%	5.35%	6.13%	43%
	CH11	83.28%	0%	16.72%	0%
	AD2	61.73%	34.13%	4.13%	71.22%
	AD4	83.05%	14.15%	2.81%	91.47%
	AD6	23.14%	0%	76.86%	8.99%
G2	CH1	74.61%	1.59%	23.8%	0%
	CH4	88.31%	0%	11.69%	0%
	CH7	85.13%	4.21%%	10.66%	58.61%
	CH9	71.68%	0%	28.32%	6.54%
	CH12	81.24%	7.74%	11.02%	4.85%
	AD1	94.47%	2.33%	3.2%	46.37%
	AD3	18.87%	2.04%	79.1%	3.26%
G3	CH2	91.86%	0.79%	7.35%	0%
	CH6	99.09%	0%	0.91%	22.09%
	AD5	99.12%	0%	0.88%	61.81%
	AD7	66.7%	5.58%	27.72%	12.34%

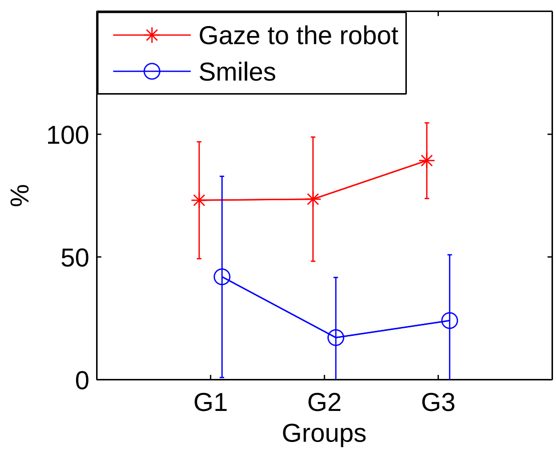


Figure 4.3: Percentage of gaze to the robot and smiles for each groups

Table 4.3: Participants' percentage of gaze behavior and smiles during the interaction with the robot

G#	**ID#**	Spoke to Nao by his/her initiative	Spoke to Nao after being encourage by his/her caregiver	Waving back gesture
G1	CH3	0	1	0
	CH5	-	-	0
	CH8	0	0	0
	CH10	0	2	0
	CH11	0	0	0
	AD2	0	0	0
	AD4	4	1	0
	AD6	0	1	0
G2	CH1	0	1	0
	CH4	2	0	0
	CH7	1	2	0
	CH9	1	1	0
	CH12	2	2	2
	AD1	4	0	2
	AD3	0	2	0
G3	CH2	1	1	1
	CH6	2	0	3
	AD5	1	0	0
	AD7	2	0	0

Table 4.4: Descriptive comments on the behavior of the participants during the interaction

G#	ID#	Comments
G1	CH3	He was really amazed by the robot, and switched his gaze to his caregiver to show his amazement
	CH5	Her caregiver was really impressed by the concentration she showed for the robot
	CH8	-
	CH10	-
	CH11	His caregiver was impressed that he looked that much at the robot.
	AD2	She was more talking and looking at the caregiver than to Nao, but was enthusiast to participate in the task
	AD4	She was really amazed by the robot and the task was corresponding to one of her routine: asking the name of the person near her
	AD6	-
G2	CH1	-
	CH4	He was really excited and impressed to see the robot (this boy asked to participate in the project)
	CH7	First, he was impressed by the robot and stand back. He finally approached it and gave it a kiss at the end of the interaction.
	CH9	She was impressed by the robot and stand back. She moved several times across the room, without giving that much attention to the robot.
	CH12	-
	AD1	She was scared by the robot when its arms were moving towards her
	AD3	It has been reported that he was more withdrawn since few weeks, and was unwilling to participate in tasks
G3	CH2	-
	CH6	He was particularly enthusiast to see the robot moving and talking to him (saying his name, waving)
	AD5	He showed some reluctance to the robot, by saying "we should put it in the garbage"
	AD7	He was reported by his caregiver by being really shy to new things

Participants showed numerous smiles, and looked at the robot multiple times. The statistical analysis only showed a relation between the participants' groups and their answer to Nao, when initiated by their own. The most visual participants showed more free speech than the other participants, and the most proprioceptive participants were the less talkative. There was no statistical evidence that the most visual participants looked more the robot during the interaction than the other participants, yet we did observe a tendency. However, the most proprioceptive participants appeared to show more smiles than the visual participants. Unfortunately, this first experiment did not permit us to validate that the behavior of the participants was linked to their proprioceptive and visual profiles. However, we have some encouraging results going in the direction of our hypothesis.

Chapter 5

Task #3 : Response to Joint Attention from a Robot

In this chapter, we explain how we designed and evaluated a Joint Attention task for our 19 participants with ND. We designed a matching game with the Nao robot. In line with our hypothesis, we observed that participants with an overreliance on proprioceptive cues and hyporeactivity to visual cues missed more the prompting and seemed to follow the prompting of the robot more slowly than individuals with an overreliance on visual cues and a hyporeactivity to proprioceptive cues.

5.1 Introduction

5.1.1 Joint Attention in ASD

Joint Attention (JA) deficit is used as a clinical sign for ASD diagnosis [Johnson et al., 2007]. JA reflects the degree to which an individual coordinates attention with a social partner toward objects, thanks to pointing and/or gaze [Mundy and Newell, 2007]. The development of JA begins in the first year of life. At approximately eight months old, an infant is able to follow his/her parents' gaze, and at 10-12 months, children begin to follow a moving target point [Johnson et al., 2007]. JA is an essential, typical, and spontaneously occurring behavior in human communication and is involved in learning. JA appears to be impaired in individuals with ASD [Johnson et al., 2007, Mundy and Newell, 2007]. Individuals with ASD tend to have impaired production and comprehension of JA behaviors. They do not use gestures or other strategies, such as finger pointing and grasping the hand of an adult, to share interest in objects or their properties, and they have impaired responses to JA initiations [Charman et al., 1997, Johnson et al., 2007, Mundy and Newell, 2007]. Intervention approaches for increasing JA have shown positive effects on social learning and development in individuals with ASD [Johnson et al., 2007, Mundy and Newell, 2007]. They are encouraged when a child is diagnosed with ASD [Johnson et al., 2007]. Several interaction therapies are used to reinforce or learn JA behaviors. They vary along the naturalistic-discrete-trial continuum: naturalistic interactions are similar to parent-child interactions, while discrete interactions use training and practice methods [Yoder et al., 2006]. For example, a study examine the effects of adult imitation on three JA behaviors of nonverbal preschoolers with autism including referential looking, gaze following and gesturing to the adult [Ezell et al., 2012]. These authors observe that adults imitating preschoolers with autism elicit JA behaviors, highlighting the value of imitation as an intervention.

5.1.2 Robots in Joint Attention Therapy for Individuals with ASD

In the field of SAR [Feil-Seifer and Mataric, 2005], robots are used as tools in socialization therapies for children with ASD to enhance social engagement, imitation, or JA skills. A common task in JA therapies with robots consists in prompting the child to look in a given direction using increasing levels of information from the robot (moving the head / moving the head and pointing with the arm) and from the target (static image / image and music / video) [Bekele et al., 2014, Anzalone et al., 2014, Warren et al., 2015, Robins et al., 2004].

In [Bekele et al., 2014], the authors evaluated the application of an adaptive robot-mediated system for JA therapy, with the Nao robot. The system was capable of both administering and automatically adjusting JA prompts to six preschool children with autism spectrum disorders and a control group formed of six TD children. They compared the answers of children to JA prompts coming from a human or a robot partner. In their setup, they used a wearable hat to detect the gaze direction of children. Initially, the subject pool was composed of 18 children (10 ASD, 8 TD). Three children with ASD were unwilling to wear the hat and were excluded from the study. In addition, one child with ASD and two TD children showed distress and were not able to participate. Results showed that children in both groups spent more time looking at the robot. They were able to achieve a high level of accuracy across trials. However, both ASD and TD children required higher levels of prompting to successfully orient within robot-administered trials. In [Anzalone et al., 2014], the authors also observed JA responses to a Nao robot or to a human partner with 16 children with ASD and a control group formed of 16 TD children. In their setup, they used RGB-D data from Kinect on the whole-body: children had to stand up during the experiment, which was three minutes long. Results showed that both children with ASD and TD children performed better with the human partner than with the robot. In [Warren et al., 2015], the authors observed JA responses to a Nao robot on six children with ASD. To track the gaze of children gaze, they used an eye-tracker system. The participant's performance was streamed to a technician, who was validating manually the child's gaze as being correct or not. Results showed that across a series of four sessions over two weeks, children improved their ability to orient to prompts administered by Nao and continued to display strong attention toward it over time. The authors did not observe if the level of impairment of children with ASD was correlated to their improvement or not. JA was also studied in the Aurora project [Robins et al., 2004], which investigates the possible use of robots in the therapy and education of children with ASD. In this study, the authors used Robota, a 45 cm high, humanoid robotic doll [Billard, 2003, Billard et al., 2007]. Four children with autism were repeatedly exposed to Robota over a period of several months, with the aim of encouraging imitation and social interaction skills. results showed that children directed their gaze and their attention towards the robot at the end of the training period than at the beginning of the training period.

Using a robot for JA therapies with individuals with ASD raises several questions: several studies showed that, compared to a human partner, a robot needs to elicit more prompting to obtain a JA response from children [Bekele et al., 2014, Anzalone et al., 2014]. However, in [Warren et al., 2015], the authors observed that JA therapy with a small group (N=6) enables children to increase their JA response. Using a robot can also permit the therapist to assess more accurately the level of prompting needed for each child. Indeed, a robot's behavior can be more easily systematically controlled than the live behavior of a human experimenter.

5.1.3 Objectives and Hypotheses

In this chapter, we present the design of a JA task designed for individuals with ND, tested with our set of participants composed of 19 children, teenagers, and adults with ND. The task is designed as a card matching-game, described in Section 5.2. Thanks to this JA task, we search to validate our hypothesis **H1**. More precisely, we hypothesize that:

- **H1'** Individuals with an overreliance on proprioceptive cues and a hyporeactivity to visual cues should less easily follow the movement of the robot than participants with hyporeactivity to proprioceptive cues and an overreliance on visual cues.

In our setup, we controlled the level of complexity of bodily signals displayed by the robot Nao (see Appendix B.1). The robot displayed different levels of prompting (Head only; Arm only; Head & Arm). We aimed to observe if the embodiment of the prompting had an impact on the groups response to JA. For the matching game, we used 15 cards divided into three subsets of visual complexities, described later in Section 5.2. Overall, individuals who have a hyporeactivity on visual cues should achieve better results than individuals who have an overreliance on visual cues [Walter, 2007]. Based on this, we believe that:

- **H1"**. An individual with a hyporeactivity to visual cues should make fewer mistakes selecting matching cards (confusing the same object in different sets).

The rest of the chapter is organized as follows: Section 5.2 describes the design and the evaluation protocol of the JA task. Section 5.3 presents the results of the interaction between the robot and the 19 participants with ND. Section 5.4 discusses the results and concludes the chapter.

5.2 Matching game with Nao

5.2.1 Design

With the help of the caregivers from our partner facilities, we designed a matching game involving a JA task. This type of game is relevant for our research goals for the following reasons: (1) it allows us to assess the perception and integration of visual cues; (2) we were able to design it to assess JA skills, and it can be tested on its capacity to reinforce JA; and (3) it can easily be adapted to different participants by using different images (for example with various levels of details).

5.2.2 Method

The experimental setup for the task is shown in Figures 5.1 and 5.2. We used Nao to conduct the JA task. It is seated on a desk. A Kinect[1] camera is placed on top of it with the help of a wooden structure built for this experiment. Two identical monitors are used. One monitor is located on the right of the robot and a second monitor is located on the left of the robot. In addition to recordings from the Kinect camera, we used two other cameras to record the interaction. The participant sits in front of the Nao robot next to the caregiver. The experimenter is behind, hidden from the participant and the caregiver, and follows the experiment procedure on a laptop. The paper cards to match are lying on the desk between the participant and Nao. The participant has to select and show to the robot the same image as the one displayed and prompted to him/her on one of the monitors.

We tested different levels of prompting. The robot prompted the participant with a sentence to look at one of the monitors (right of left) and by a movement of its (1) Head only, (2) Arm only, or (3) Head & Arm. The movements were selected in a random order with five movements in the Head only condition, five movements in the Arm only condition, and five movements in the Head & Arm condition. The right and left monitors were also selected randomly. We used 15 cards for the matching game, divided into three subsets of visual graphical complexity.

[1]urlhttps://developer.microsoft.com/en-us/windows/kinect

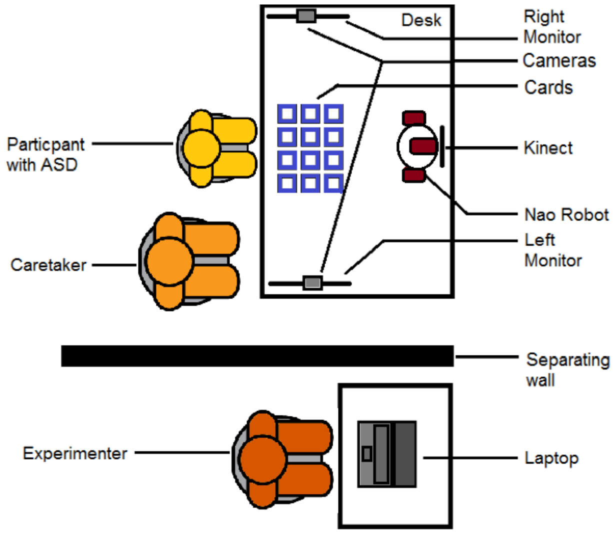


Figure 5.1: Setup of the Joint Attention task experiment

Images in the first subset are simple drawings (Figure 5.3). Images in the second subset are more complex drawings (Figure 5.4) in the sense that they have a higher level of graphical details. Images in the third subset are photos (Figure 5.5). Each subset is composed of images of the same set of objects: a tree, a bike, a car, a house, and a teddy bear. The cards are composed of the images printed on white paper on squares with sides 6 cm long, glued on a blue squared cardboard with sides 10 cm long. We drew the images from subset one and selected public domain images for the subsets two and three. We chose not to select the images from databases specialized for individual with ND for two reasons: (1) they are costly; (2) available databases do not provide different images of different graphical complexities for the same concept.

Before starting the interaction, the way to show cards to Nao was explained to the participants. We taught children with ND how to show the cards with an example set formed by two cards that were not used in the experiment (rabbit and dog drawings). This part was performed several times until the child with ND was able to correctly show the cards to the robot.

The interaction steps are as follows: (1) Greetings: Nao says, "Hello, I am happy to see you, *participant name*!", while looking at the participant; (2) Game: Matching game, see Figure 5.6; and (3) Goodbye: Nao says goodbye to the participant ("Goodbye, *participant name*, see you soon!"). The caregiver is instructed not to interact with the participant during the entire session to avoid inducing JA.

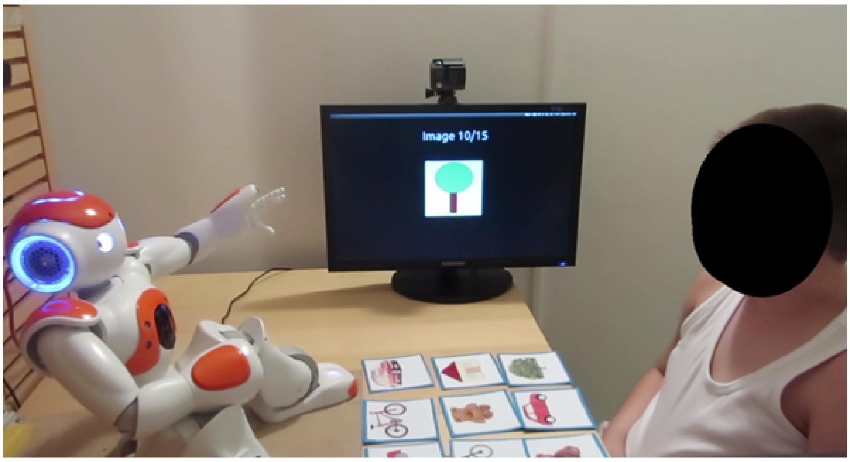


Figure 5.2: Matching game with the robot Nao

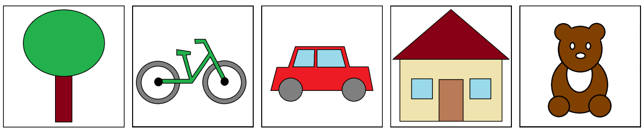


Figure 5.3: Images of the first subset (very simple sketches with a small number of details). From left to right: Tree1; Bike1; Car1; House1; Bear1

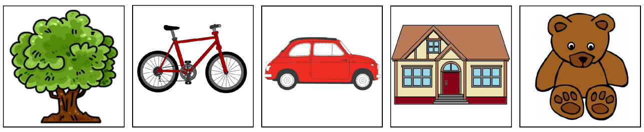


Figure 5.4: Images of the second subset (sketches closer to the reality with more details than in the first subset). From left to right: Tree2; Bike2; Car2; House2; Bear2

Figure 5.5: Images of the third subset (real images with real objects found in our daily life). From left to right: Tree3; Bike3; Car3; House3; Bear3

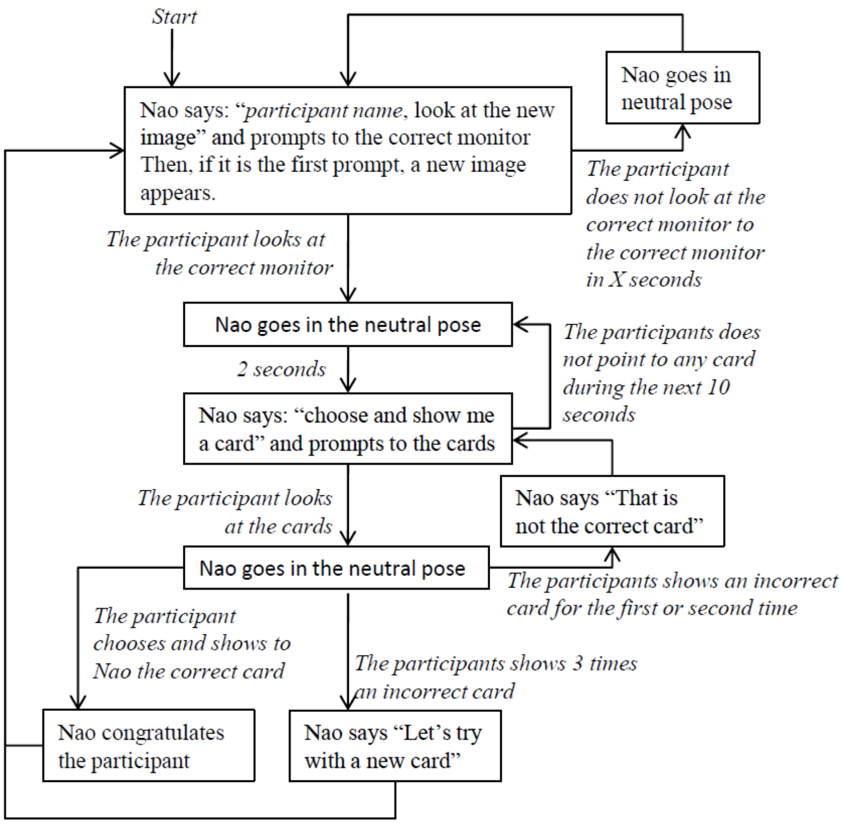


Figure 5.6: Diagram of interaction with the Nao robot

5.2.3 System Architecture

The system to lead the interaction is divided into three main components: (1) the multiple head pose tracker, (2) the NAOqi[2] object recognizer, and (3) the main game software controlling the interaction between NAO and the user(s). Each software component runs independently and communicates with the other resources using ROS[3] topic publishing/subscribing. The multiple head pose tracker and the main game software run on a regular computer, while the NAOqi object recognizer runs on the NAO robot itself. The software running on the computer has been developed in C++ and can be compiled to run on any operating system that supports ROS.

The Multiple Head Pose Tracker detects, identifies, and tracks the persons standing in front of the camera, estimating their head pose as well as their position in real world coordinates. The system can be configured to track virtually any number of persons. The software uses the RGB and depth images acquired from a Kinect camera (using the openni ROS package),

[2] http://doc.aldebaran.com/2-1/naoqi/
[3] https://github.com/ros-drivers/openni2_tracker

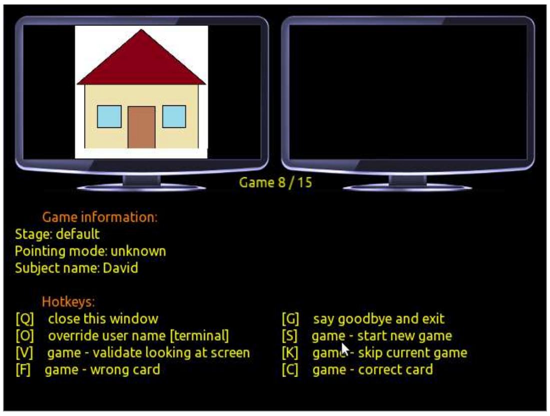


Figure 5.7: Screenshot of the control panel that permit the experimenter to follow the interaction and to override manually the behavior of the robot.

but it can also be used with a standard webcam (using the gscam ROS package), with limited functionality. We first use the basic OpenCV HaarCascade detector to detect regions of interest of the heads, then we used Saragih's facetracker [Saragih et al., 2009] on the detected regions to determined the orientation of the heads.

We developed a manually override to permit the experimenter to take the hand on the program, if needed (for example if the face of the individual with ND is not detected because of his/her position or because he/she hides a part of his/her face with hands).

The experimenter leads and follows the interaction thanks to the Kinect camera output and a control panel we designed (see Figure 5.7).

The interaction was partly simulated using a WOZ protocol [Kelley, 1983] involving an experimenter in the interaction loop. During such interactions, the behaviors of participants with ND would be difficult to track automatically. They cover their faces with their hands or the cards, or they look straight down which makes the automatic video detection of their face and gaze challenging.

5.2.4 Mesures

With this experiment, we could test our hypotheses. We wanted to verify that the behavior of our participants was linked to their proprioceptive and visual profiles (see Chapter 2). In order to analyze the interactive behaviors of the participants, we collected additional informa-

Table 5.1: Inter-Coder Reliability

	Inter-Coder Reliability	Percentage of data coded by the two coders
Answer to prompting	$\kappa = 1$	27.10%
Gaze to the correct monitor	$\kappa = 0.824$	26.80%
TRJA	$ICC = 0.92$	27.10%
Greetings to robot	$\kappa = 0.88$	42.10%

tion to assess their social skills inspiring from the Early Social Communication Scales (ESCS) [Seibert and Hogan, 1982] and existing Human-Robot Interaction metrics [Begum et al., 2015]. The measures that we computed were the following:

- *Response to JA*: correct following of the pointing gestures and gaze displayed by the robot. We observe if the participant missed or looked at the wrong monitor when prompted.
- *Time of Response of JA* (TRJA): the time between the beginning of the movement of the robot and the moment the participant's eyes hit the screen.
- *Mistakes*: we recorded the answers from the participants during the matching game and collected their mistakes and correct responses.
- *Emotional ans social behaviors*: we recorded the amount of time where the participants were smiling and the emotional ans social behaviors of the participants during the interaction with the robot.

A first coder annotated all the 19 videos. A second coder, unaware of the hypotheses of the setup, annotated parts of the videos (Table 5.1). The Cohen Kappa Coefficient (κ) and Intraclass Correlation Coefficient (ICC) were computed to assess the inter-coder reliability.

5.2.5 Data Analysis

We performed a repeated-measure ANOVA on participants' TRJA with two factors considered for the analysis: channels *conditions* with 3 categories (Head Only; Arm Only; Head & Arm),

Table 5.2: Response to Joint Attention (JA) and mistakes made by the participants during the game

G#	ID#	Missed prompts	Gaze to the wrong monitor	Mistakes
G1	CH3	1	2	None
	CH5	1	0	No data
	CH8	6	0	Bear2 instead of Bear1; Bear3 instead of Bear1; Bike2 instead of Bike3; Car2 instead of Car1
	CH11	2	2	No data
	AD2	0	2	None
	AD4	0	1	None
	AD6	4	1	Bear3 instead of Bear2
G2	CH1	0	0	None
	CH4	0	0	None
	CH7	0	0	None
	CH9	2	0	Bike2 instead of Bike3
	CH12	0	1	None
	AD1	0	4	None
	AD3	0	0	None
G3	CH2	0	0	Tree3 instead of Tree2
	CH6	0	0	None
	AD5	0	3	Bear2 instead of Bear1; Bear3 instead of Bear1; Car2 instead of Car3
	AD7	0	0	None

and *repetitions* (repetition 1 to 5). We performed Fisher's Exact Test on the number of participants that missed prompt, gazed to the wrong monitor and did mistakes. The significance threshold was set to $p < 0.05$. We used Statistica version 13 to perform the analyses.

5.3 Results

Tables 5.2 and 5.3 describe the participants' behaviors during the interaction with the robot, and Table 5.4 displays the mean TRJA of each participants during the experiment.

Table 5.3: Emotional ans social behaviors expressed by participants during the interaction with Nao.

G#	ID#	Social Interaction	Negative/Neutral Emotional Behavior	Smiles (%)
G1	CH3	-	Not focused, turned his attention towards the caregiver	11.91%
	CH5	-	Boredom (at the end of the interaction)	0%
	CH8	-	Tired	55.33%
	CH10	Answered to hello and goodbye	-	2.29%
	CH11	-	-	0%
	AD2	-	Not focused, turned her attention towards the caregiver	52.4%
	AD4	Talked to the robot, Answered to goodbye	-	72.76%
	AD6	-	Showed signs of being annoyed, and refused to perform the task for the first six images of the experiment	0%
G2	CH1	-	-	1.75%
	CH4	Answered to goodbye	-	11.93%
	CH7	Answered to hello and goodbye, kissed the robot when leaving the room	-	25.57%
	CH9	Smiled when greeted, answered to goodbye	-	61.58%
	CH12	Said hello to the robot before the robot said hello, Answered to goodbye	-	4.16%
	AD1	Talked to the robot, Answered to goodbye	-	39.82%
	AD3	-	No emotional behavior	0%
G3	CH2	-	No emotional behavior	52.4%
	CH6	-	-	11.38%
	AD5	-	Showed signs of boredom and annoyance during the second half of the experiment	37%
	AD7	-	-	6.17%

Table 5.4: Mean Time of Response to Joint Attention (TRJA) The mean TRJA for each condition (i.e., Head only, Arm only and Head & Arm) and for the whole interaction (total) are presented.

G#	**ID#**	**Head** (s)	**Arm** (s)	**Head & Arm** (s)	**Total** (s)
G1	CH3	5.42 ± 3.29	2.78 ± 0.51	5.43 ± 4.93	4.48 ± 3.28
	CH5	2.53 ± 0.40	2.29 ± 0.42	3.59 ± 1.47	2.88 ± 1.08
	CH8	4.05 ± 1.71	9.82 ± 10.2	5.43 ± 3.91	5.62 ± 4.66
	CH10	2.5 ± 0.15	2.93 ± 0.96	2.53 ± 0.22	2.65 ± 0.57
	CH11	12.13 ± 4.72	5.8 ± 4.56	13.2 ± 12.3	9.77 ± 7.33
	AD2	4.19 ± 2.09	2.53 ± 0.36	2.64 ± 0.38	3.12 ± 1.39
	AD4	2.88 ± 0.31	2.92 ± 0.83	4.06 ± 2.08	3.29 ± 1.33
	AD6	2.81 ± 0.27	2.66 ± 1.07	2.11 ± 0.84	2.50 ± 0.82
G2	CH1	2.23 ± 0.47	2.93 ± 0.52	4.35 ± 3.54	3.17 ± 2.13
	CH4	1.23 ± 0.24	1.55 ± 0.30	1.39 ± 0.36	1.39 ± 0.31
	CH7	1.13 ± 0.6	1.80 ± 0.71	1.67 ± 0.69	1.53 ± 0.69
	CH9	2.65 ± 0.4	3.03 ± 1.51	2.74 ± 0.43	2.79 ± 0.81
	CH12	2.30 ± 0.77	2.42 ± 0.15	2.19 ± 0.57	2.30 ± 0.53
	AD1	2.92 ± 0.71	2.22 ± 0.53	2.49 ± 0.84	2.54 ± 0.72
	AD3	4.73 ± 1.30	3.21 ± 1.87	2.83 ± 0.30	3.59 ± 1.49
G3	CH2	2.07 ± 0.46	2.10 ± 0.42	1.59 ± 0.67	1.92 ± 0.55
	CH6	3.29 ± 2.47	3.51 ± 0.88	3.43 ± 1.81	3.41 ± 1.70
	AD5	2.21 ± 0.80	4.55 ± 2.13	2.60 ± 0.70	3.12 ± 1.66
	AD7	1.71 ± 0.43	2.29 ± 0.63	2.22 ± 1.04	2.07 ± 0.74

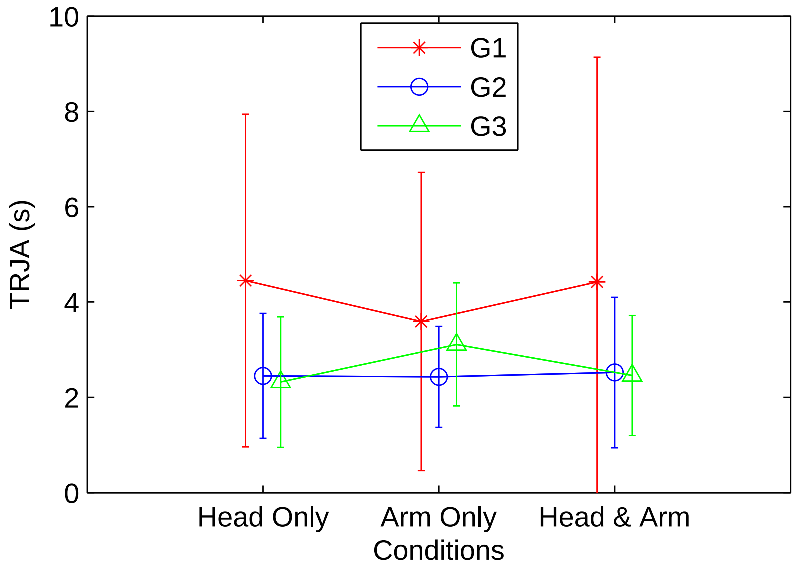


Figure 5.8: TRJA for each group for the whole experiment, Head Only condition, Arm Only condition and Head & Arm condition

5.3.1 Response to Joint Attention

As shown in Table 5.2, participants from groups G2 and G3 responded well to JA when prompted by the robot. Except for CH9 (G2), none of them missed a prompt from the robot and except for CH12 (G2), AD1 (G2), and AD5 (G3), none of them looked at the wrong monitor. Participants from group G1 showed more difficulties to follow the robot's movements compared to participants from groups G2 and G3. Five out of seven participants from group G1 missed at least one prompt from the robot, and six of them looked at the wrong monitor. The number of participants that missed at least one prompt was significant among the groups ($p < 0.05$). There was more participants from group G1 that missed at least one prompt (five on seven), than participants from group G2 (one on seven), and participants from group G3 did not miss any prompt during the interaction with the robot (see Table 5.2).

5.3.2 Time to Response to Joint Attention

We found no significant main or interaction effects of the groups, age, channels conditions or repetitions on the TRJA of the participants. Participants did not improve their TRJA with the repetitions (Figure 5.8).

5.3.3 Mistakes

Some participants from all groups (CH8 (G1), CH10 (G1), AD6 (G1), CH9 (G2), CH2 (G3), and AD5 (G3)) made mistakes when choosing the cards.

5.3.4 Emotional ans social behaviors

We found no significant main or interaction effects of the groups, age, channels conditions or repetitions on the smiles of the participants. Participants from all groups displayed positive behaviors (Table 5.3). Participants from group G2 showed more social behaviors (i.e., answer to greetings) than the participants from the other groups. Participants from group G1 showed more negative behaviors (e.g., signs of boredom, of being annoyed, etc.) than the participants from the other groups.

5.4 Conclusion

In this chapter, we described the design of a JA interaction task (a matching game) with Nao for individuals with ND. In Chapter 2, we had defined individual profiles by considering the integration of proprioceptive and visual cues of each individual.

In this experiment, we did not found statistical evidences that proprioceptive and visual profiles had an effect on the time of response to JA. This may be caused by the small number of participants, by the low number of repetitions and because the experiment was not repeated over time. The absence of impact of the embodiment on the TRJA may not be surprising as in [Cooney et al., 2015], the authors enlighten that, in TD population, there are evidence that a common neural mechanism process directional social cues (i.e., eyes, head, body and arms directional cues), and this process may be blind to the stimulus category. Participant from group G1 missed more prompt than the other participants, and participants form group G3 did not miss a single prompt from the robot. These results goes in the direction of the hypothesis **H1'**. Mistakes were made by four participants from the three groups. These results nevertheless do not permit us to validate **H1"**. Overall, participants from all the groups showed positive emotional behaviors such as smiles. But we observed that the matching game was difficult for participants with low cognition. We observed signs of boredom at the end of the sessions for some participants.

In [Haswell et al., 2009] and [Izawa et al., 2012], the authors observed that the greater the reliance on proprioception, the greater the child's impairments in social function and imitation. These authors suggest a greater-than-normal dependence on cortical regions in which movements are represented by intrinsic coordinates of motion (M1 and somatosensory cortex) and a less-than-normal dependence on regions in which movements are represented by extrinsic coordinates (premotor and posterior parietal) in children with ASD. The parietal and superior temporal cortices support the response to JA development in infancy [Mundy and Newell, 2007]. They serve aspects of representational development, imitation, and the perception of the eye and head orientations of others, as well as the perception of spatial relations between the self, others, and the environment. Our results showed that participants with a greater reliance on proprioception display a slower response to JA, confirming the results from another studies [Haswell et al., 2009, Izawa et al., 2012].

In [Baranek et al., 2013], the authors observed the answers to social and non-social visual, audio, and touch sensory stimuli in children with ASD. They observed that more the children showed answers to both the stimuli more they showed mental ability, and less they showed answers, the less they were skilled in JA. As in this study, we observed that hyporeactivity to visual cues lead to slower response to JA in children with ND.

There are also limitations on our work that should be outlined. We used a WOZ setup because of the difficulty of tracking participants' faces. We did not want to use a wearable hat [Bekele et al., 2014] or other invasive devices because we knew that most of our participants would not tolerate wearing them for an entire experiment. Using eyetracking technology would enable to conduct fine grained analyses of gazing behaviors, which would provide more information about JA interactive behaviors [Noris et al., 2012].

Chapter 6

Task #4 : Imitation of Movements with a Robot and a Human

In this chapter, we describe the last task that we defined for testing our hypothesis. We selected an imitation task because of its importance in social interaction and its relevance for participants with ND. The experimental protocol was composed of seven imitation sessions over eight weeks. During some sessions, the participant was imitated or imitated a human. In other sessions, the participant was imitated or imitated a Nao robot. As expected, we observed that children with an overreliance on proprioceptive cues and hyporeactivity to visual cues had more difficulty imitating the robot than the other children. Moreover, for most of the children, the repeated sessions had a positive impact in terms of interaction behaviors (gaze to the partner, imitations) toward a human partner.

6.1 Introduction

6.1.1 Imitation and ASD

Imitation is a very important element in child development. Imitation is described as one of the precursors of social cognition. As explain in [Nadel, 2014], imitation has different forms: (1) immediate imitation (i.e., doing the same thing as another person at the same time as them);

(2) delayed imitation (i.e., doing the same but after a delay); and (3) deferred imitation (i.e., doing the same but much later and without the person); and have different social consequences. Immediate imitation has a communication function, while delayed imitation has a social interaction function (e.g., as in a "do like me" game) and deferred imitation has a learning function [Nadel, 2014].

In [Nadel, 2002], synchronous imitation is shown to be a way for preverbal children to communicate, and there is a predominance of the use of imitation during social exchanges at approximately 18 months of age with a peak at 30 months. Imitation behaviors occur very early in children. In [Kugiumutzakis, 1998], the author showed that imitation could be observed in neonates who were less than 45 minutes old. In individuals with ASD, studies have observed that there were deficiencies in imitation [Charman et al., 1997, Rogers and Pennington, 1991, Williams et al., 2004]. In [Vanvuchelen et al., 2007] the authors discussed the causes of this impairment and noted that imitation deficiencies could occur due to motor or cognitive deficiencies: in ASD, the impairment in imitation seems to be caused by motor perception impairment instead of cognitive perception impairment. However, [Charman and Baron-Cohen, 1994] argued that imitation impairment is preceded by the alteration in cognitive development. In addition, there is a substantial variability in imitation deficiencies in children with ASD. Some children with ASD can learn to imitate or can imitate spontaneously, while others have larger impairments. However, the dynamics of imitation (i.e., turn taking, initiation of new behaviors, and adaptation to other's behavior) are altered in the large majority of individuals with ASD [Tomasello, 2009, Carpenter et al., 2005, Nadel and Butterworth, 1999]. A positive impact of imitation-based therapies on social skills was observed in children with ASD [Dawson and Adams, 1984, Dawson et al., 1990, Nadel-Brulfert and Baudonniere, 1982, Nadel, 1986, Field et al., 2001]. In [Dawson and Adams, 1984], the authors found that children with ASD and a low level of imitative ability showed more social response and eye contact, and they played with toys in a less preservative manner when the experimenter was imitating their behaviors. Again, in [Field et al., 2001], the authors showed that when children with ASD were imitated by the adults, they looked more at the partner, smiled more, engaged more in reciprocal play, sat next to and close to the adult and touched the adult more than when only adults were only playing with them. In [Nadel, 2004], the authors found that there was a significant correlation between the level of imitation and level of imitation recognition in children with ASD. Some children with ASD recognized they were imitated. Those children showed different strategies to test the intentionality of their partners, such as: changing the rhythm of activity, changing the object used, and stopping the current action while gazing at the partner. Some children did not test their partners, but their attention increased and they showed more positive affect towards the partner. This may indicate that the children were aware they were being imitated without understanding the intentionality of the partner.

6.1.2 Robots in Imitation Therapy for Individuals with ASD

Numerous studies observed the effects of a robot partner in imitation therapy for children with ASD [Robins et al., 2004, Duquette et al., 2008, Tapus et al., 2012, Ranatunga et al., 2013, Taheri et al., 2015, Conti et al., 2015]. In [Klin et al., 2009], it was observed that children with ASD had impaired detection of biological motion compared to non-biological motion. In [Diehl et al., 2012], the authors suggested that a task involving imitating robots could have a positive impact on individuals with ASD compared to a task involving imitating humans. The authors in [Robins et al., 2004] observed in a longitudinal study with the Robota robot that social behaviors, such as eye gaze, touch, and imitation increased over time in children with ASD. In [Duquette et al., 2008], the authors observed that the imitation of body movements and of familiar actions were more frequent when the children with ASD were paired with a human partner than with a robot partner. However, children with ASD who were paired with a robot showed increased shared attention and imitated facial expressions more than the children paired with a human partner. The authors in [Tapus et al., 2012] observed the social engagement (gaze towards the robot, smiles, initiation of movements) of four children with ASD and investigated whether they were more engaged with a Nao robot or with a human partner. There was a high variability in the reactions to the Nao robot. The robot did not impact the social engagement of two of the children. The other two children gazed more at the robot and displayed more smiles or laughter with the robot compared to the human. One performed better in the imitation task with the Nao robot than with the human partner.

In the above studies, the authors described that SAR for training and learning imitation with individuals with ASD was effective on their social skills. Training the skills across several session is important for learning effect. Indeed, individuals with ASD shows difficulties to learn new skills and to generalize them for new situation and stimuli [Froehlich et al., 2012]. In addition, SAR protocol for individuals with ASD, it is important to verify if the skills learned with the robot are enhance towards human partner [Feil-Seifer and Mataric, 2005, Tapus et al., 2007]. To verify the generalization towards a human partner is an important step in a robot-base protocol.

6.1.3 Objectives and Hypotheses

In this chapter, we present the design of an robot-based imitation experimental protocol for children with ND. In this study, we observed the effect of repeated sessions on the participants. We observed in the Chapters 3, 4 and 5 that a link exists between the profiles defined in Chapter 2 and the interaction skills of the participants. In this task, we wanted also to observe the effect of the profiles on learning a skill and generalization towards a human partner. We worked only with the children participants of our subjects pool ($N = 12$). During the sessions, the children were instructed to imitate a partner, and were imitated by the partner. With this imitation task, we aimed to validate the relevance of the proprioceptive and visual individual profiles we defined in the Chapter 2. We sought to validate our hypothesis **H1**. More precisely, we hypothesized that:

An individual with an overreliance on proprioceptive cues and a hyporeactivity to visual cues will:

- **(H1A)** struggle more to imitate his/her partner;
- **(H1B)** struggle more initiating movement for his/her partner to imitate him/her;
- **(H1C)** display slower improvements in his/her imitation skills across the sessions;

than an individual relying on visual cues.

The remainder of the chapter is organized as follows: Section 6.2 describes the design and evaluation protocol of the imitation task. Section 6.3 presents the results of the interaction between the robot and twelve children with ND. Section 6.4 presents the chapter's conclusions and a detailed discussion.

6.2 Design of the Imitation Task

6.2.1 Content of the Interaction

With the help of the caregivers from our partner facilities, we designed an imitation task with Nao (see Appendix B.1). We designed it as an ABAB therapy, in which the participants' partner in the experiment was a human in A session, and Nao in B sessions. The children knew the human partner but had never worked with her in therapeutic sessions. As all of the children met the robot before, and knew the experimenter, there were no expected novelty effects in any of the partners (human or robot). The sessions were organized identically for all participants, as described in Table 6.1. There was seven sessions for each participant. Each session had a code name: H# for a session with a human partner, and R# for a session with a Nao robot as partner, with # from 0 to 6 for each number of sessions. The first and fifth sessions (H0 and H5) were with the human partner. The other sessions (R1, R2, R3, R4, and R6) were with the robot. The sessions H0 and R1 were pre-test sessions. Sessions R2 to R4 were the repeated sessions. After session R4, we observed a break of one week. Then, the 6th and 7th sessions (H5 and R6) were post-test sessions, enabling us to observe if there were generalization effects on the imitation skills of the participants.

The sessions were held in a reserved room in each of the care centers. When the session was with the robot, the robot was placed on a table. An Xtion RGB-D camera was fixed in front of the child, and 2 cameras recorded the sessions. The child could sit or stand in front of the partner. Figure 6.1 shows the setup for both partners. There are two phases in each session: (a) the child imitates the partner (human/robot) and (b) the partner (human/robot) imitates the

Table 6.1: Planning of the experimental protocol. Sessions are coded as follow: H# for a session with a human partner, and R# for a session with the Nao robot as partner, with # ranging from 0 to 6 for each number of sessions.

Week	Session code	Session partner
Week 0	H0	Human
Weeks 1 to 4	R1-R4	Robot
Week 5	Break	-
Week 6	H5	Human
Week 7	R6	Robot

child.

Phase (a): the child imitates the partner

In this phase, the child has to imitate the arms movements of the partner (human or robot). Several imitation tasks have been proposed in other studies for example involving objects [Nadel, 2002]. We decided to select imitation of simple movements since these motor tasks may be impacted by visual and proprioceptive preferences. The protocol was design with the help of the caregivers from the care-facilities in regards of their work with the children with ND.

To keep the attention of the participants and make the interaction meaningful, we designed it as a playful task. The partner was initiating arms movements on a song, as she/it was dancing. As we work with children with a wide range of cognition levels and motor skills, the songs lacked lyrics and the gestures were limited to the arms. Each movement is a combination of the left and right arms (L and R, respectively) taking one of the following five positions: D: down; U: up; F: in front; P: in ψ; and T: in T form, see Figure 6.2. The movements to imitate were randomly ordered to avoid a learning effect between the repetitions. For the human partner, the movements were chosen in real time, and for the robot partner, the list of movements was computed at the beginning of the session.

We used three songs, with the same rhythms and durations (1 min 30 s). During a session, two of the three songs were used. In this way, we hoped to keep the children interested, as well as to be able to alternate the songs between the sessions. A song was played a maximum of

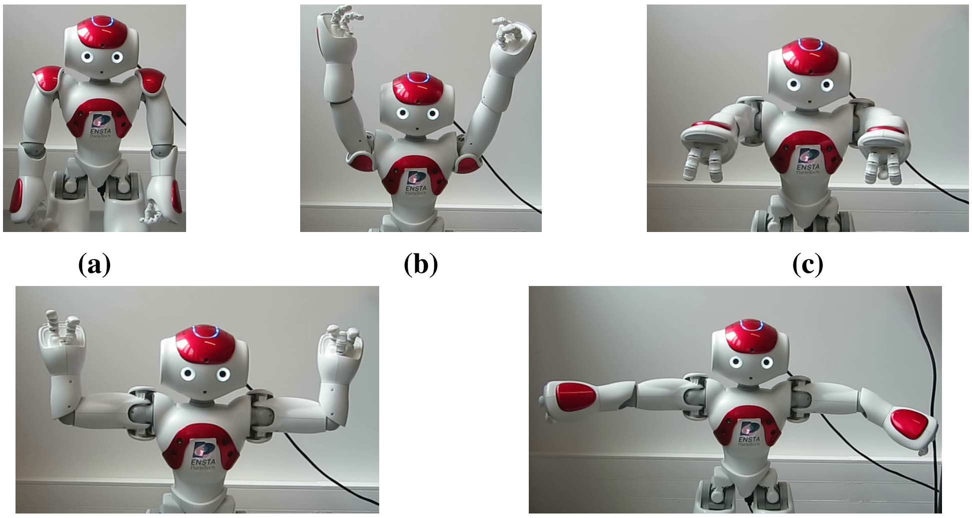
(a) with the robot partner

(b) with the human partner

Figure 6.1: Pictures of the setup

two times in a session. In ASD, unusual responses to audio sensory stimuli have been observed [Rogers and Ozonoff, 2005]. To prevent unexpected behaviors toward the songs, we asked to the caregiver for each child if they tolerated and liked music, and we had the possibility to lower or turn off the music if needed. Overall, the children liked the songs and we did not need to lower or cut the volume. The child imitated the partner a maximum of four times by session (in total, six minutes of imitation). Participants were able to take short breaks, if needed, between the dances.

We wished to propose sessions that adapt to each participant, stimulating and evolving in time. To avoid boredom across the sessions, the number of arm movements increased throughout the sessions in an adaptive manner i.e., related to 80% of movements correctly imitated (see Table 6.2 to see the movements that were imitated during a specific session). The number of correct movement was annotated by the experimenter while the child was performing the imitation. The time between the movements was modulated by the performance of the child during a dance within a session. In a session, if a participant performed 80% of the movements in two consecutive dances, the time between two movements decreased in the next dance. If the child was unable to imitate 80% of the movements the time between two movements was unchanged. The time between two movements was 8 seconds and could decrease to 6, 4, and 2 seconds, depending on the participant's performance. The time between two movements were discussed with the caregivers. The first sessions started with a movement every 8 seconds, to enable the caregivers to have enough time to correct and encourage the children to imitate the partner. It was also short enough not to lose the attention of the children. The decreases were chosen to gradually increase the attention of the children towards the partner. In sessions H0 and R1, the partner produced a movement every 8 seconds. At the beginning of the R2, R3 and R4 sessions, the time between two movements was set as the one reached at the last dance of the previous session. In the H5 and R6 sessions, the partner started the session with the frequency reached in the session R4. If a participant reached the level of a movement every 2 seconds, the first dance was reset to a movement every 4 seconds in the next session. Indeed, to follow this frequency of a movement every 2 seconds was quite demanding to maintain for 4 dances.



(a) (b) (c) (d) (e)

Figure 6.2: Example of the arm positions for Nao: a: LDRD; b: LURU; c: LFRF; d: LPRP; and e: LTRT

Finally, if the participant imitated 80% of the movements, the partner cheered the child at the end of an imitation. During this phase, the caregiver was standing next to the child and was asked to help him/her understand the imitation task, encourage the child to imitate the partner and correct his/her arms position, if needed.

Phase (b): the partner imitates the child

In this phase, the partner imitated the gross arm movements of the child. The child was told to perform slow large arms movements. A song of 1 m 30 s, different from the ones of Phase (a) was played during this phase. It permitted to do a link between the Phases (a) and (b). In addition, we hoped it would encourage the child to perform movements, and it gave a time limit to perform this task. In the care facilities we work with, the children work mostly with a timer. During this phase, the caregiver remained with the child and tried to encourage him/her to perform arms movements.

Observed behaviors

The sessions were video recorded and the behaviors of the participants were manually annotated. In both phases, we computed the participants' *percentages of gaze* towards the partner

Table 6.2: Arm positions for the partner. Each movement is a combination of the left arm and the right arm (L and R, respectively) taking one of the following five positions: D: down; U: up; F: in front; P: in ψ; T: in T form

Session code	Movements
H0-N1	M1 = LDRD + LURU + LFRF + LPRP + LTRT
N2	M2 = M1 + LURP + LPRU + LTRF + LFRT
N3	M3 = M2 + LURD + LDRU + LTRP + LPRT
N4-H5-N6	M4 = M3 + LFRP + LPRF + LTRU + LURT

(human/robot) and *percentages of smiles* during the interaction. In addition, in Phase (a), the *imitation score* was assessed; in Phase (b), the *free initiations* were noted. Imitation scores were annotated as follows: "0" if there was no movement, if there was a movement that appeared to be random or not connected to the movement of the partner or incorrect; and "1" if the movement was correctly imitated. Then, we computed the percentage of the movements that was correctly imitated across the whole session. As described in Section 6.2.1, the number of movements increased if the child correctly imitated the partner. We interpreted the *number of movements* the children correctly imitated as an indicator of the improvement and success of a child to imitate the partner. Free initiation is the percentage of the time during which the child initiates movements for the partner to imitate him/her during Phase (b).

6.2.2 System Design

Phase (a)

We used the API Naoqi[1] provided by Softbank Robotics to program Nao's movements. The experimenter launched the program when the child was ready and in front of the robot. The song and gestures started together. The experimenter could stop the song and gestures at any moment.

[1] http://doc.aldebaran.com/2-1/naoqi/

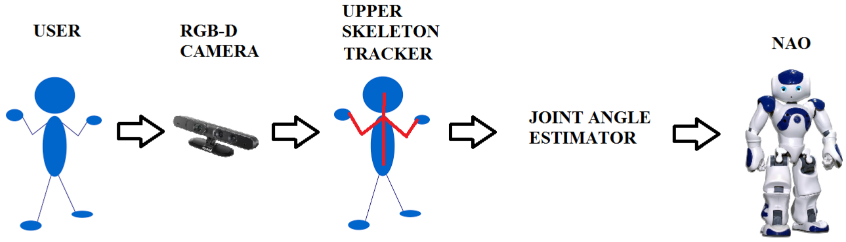


Figure 6.3: Schematics of the imitation system

Phase (b)

We used a RGB-D camera (Xtion Asus[2]) and the ROS package Openni2_tracker[3] to track the participants' upper body. The tracker estimates the skeleton of the user, a list of body joints. These joints are translated into the motor commands for the Nao robot through a joint angle estimator and Naoqi. The joint angle estimator converts the shoulder and elbow angles from the participant into scaled angles for Nao. See in Figure 6.3 a schematic of the system. The participants could stay seated and no calibration-pose was needed to start the imitation by the robot. A similar system has been used successfully in previous studies [Tapus et al., 2012].

6.2.3 Participants

In this experiment, we worked only with the children participants ($N = 12$; 11.7 ± 2.6 years old) in our subject pool presented in Section 1.3.2. In Table 6.3, we provide a reminder of each children participating to the imitation experimental protocol and the corresponding groups as defined in Section 2.2. A child, CH2 (G3), was very agitated for the final four sessions (out of seven sessions) due to personal issues. We did not include him in the statistical analysis for this task because his behavior changed significantly. As for all the other participants, we still describe his behavior.

In total, there were five children in group G1, five in group G2, and one in group G3. When forming the groups (see Chapter 2), we observed that the participants from G2 and G3 used more visual cues than participants from group G1. Because the group G3 had only one

[2]https://www.asus.com/fr/3D-Sensor/Xtion_PRO/
[3]https://github.com/ros-drivers/openni2_tracker

Table 6.3: Children's information

G#	ID#	Gender	Age	Comments
G1	CH3	M	12	-
	CH5	F	11	Low level of cognition; Non-verbal; West Syndrome (uncommon to rare epileptic disorder).
	CH8	M	15	-
	CH10	M	10	-
	CH11	M	17	Low level of cognition
G2	CH1	M	11	-
	CH4	M	13	High level of cognition. Asked to be part of the program to meet the Nao robot.
	CH7	M	8	-
	CH9	F	12	-
	CH12	M	9	-
G3	CH2	M	9	Suffers of echolalia
	CH6	M	13	-

participant, and by noting that participants from groups G2 and G3 rely more on visual cues and less to proprioceptive cues than participants from G1, we merged the G2 and G3 groups. Therefore, the new group composition for the analysis of the imitation task is as follows:

- Group G1, with the following children: CH3; CH5; CH8; CH10 and CH11;
- Group G2, with the following children: CH1; CH4; CH7; CH9; CH12 and CH6 (previously in G3);

6.2.4 Data Analysis

In total, each participant performed 7 sessions of approximately 10 minutes, interacting with the partner (human and robot), resulting in 84 videos. A first coder annotated all videos and a second coder, unaware of the hypotheses of the setup, annotated 20% of the videos, which were randomly chosen. An intraclass correlation coefficient (ICC) was computed to assess the inter-coder reliability. The ICC was 0.86, which shows good inter-coder reliability. We observed the effect of the groups on the participants' gaze direction, imitation scores, free initiation, and smiles as explained in Section 6.2.1. We performed a repeated-measure ANOVA on participants' behaviors. Post-hoc analyses were performed with Tuckey's HSD test. We performed correlations between the behaviors and the sessions to observe any learning effect of the experimental setup.

6.3 Results

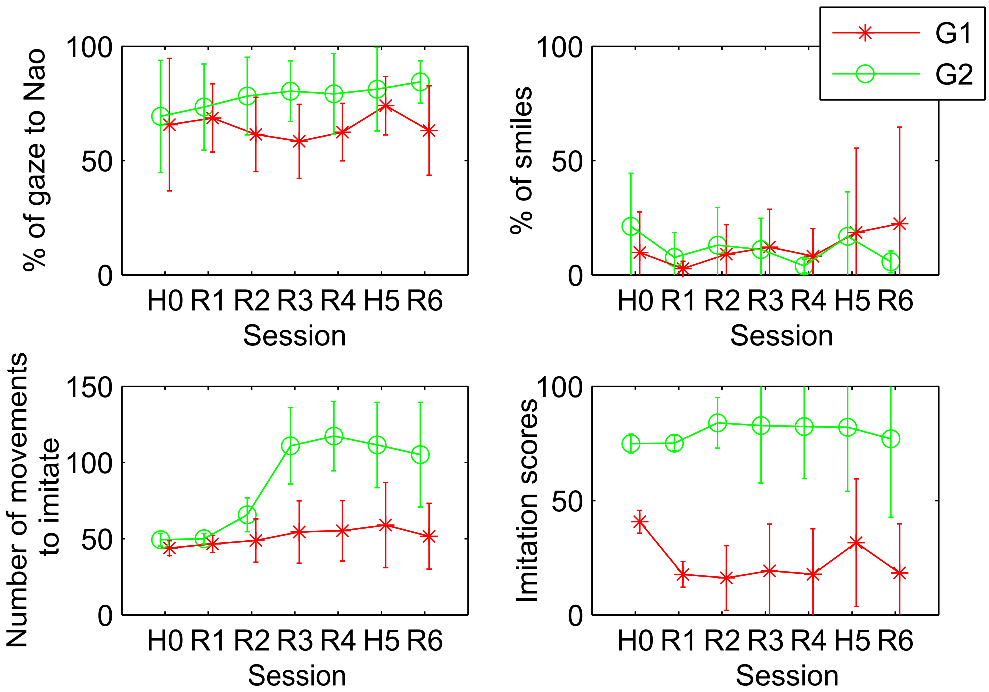
All 12 children performed all seven sessions of the experiment. They were enthusiastic to participate in the sessions, even the children who did not perform well on the task. A child, CH2 (G3), was very agitated and perturbed for the final four sessions (out of seven sessions) due to personal issues. We did not include him in the statistical analysis of the proprioceptive and visual profiles because his behavior changed significantly. As for all the other participants, we still describe his behavior.

6.3.1 Phase (a)

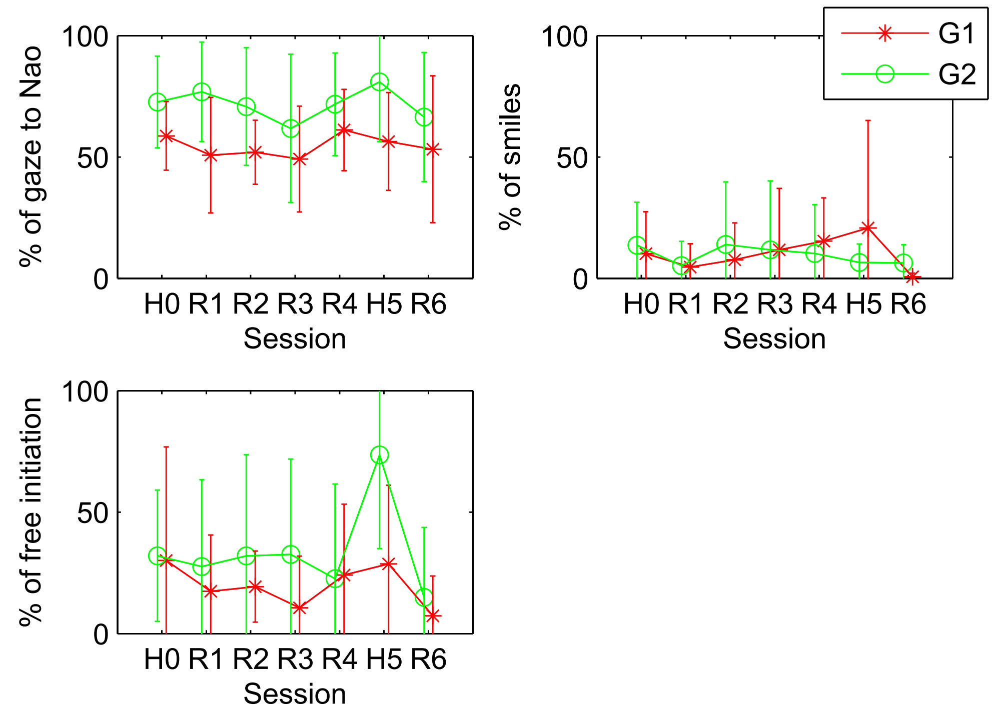
In this phase, children had to imitate the partner. Figure 6.4a show the observed behaviors for the G1 and G2 groups in Phase (a).

Gaze to the partner We observed a weak significant main effect of the groups on the gaze direction of the children during Phase (a) ($F(1, 9) = 4.21$; $0.05 < p < 0.1$). Participants from G1 looked globally less to the partner than the participants from group G2 (G1: $M = 64.83$, $SD = 17.13$; G2: $M = 78.02$, $SD = 16.89$). We observe in Figure 6.4a an increase in the mean gaze to the partner, through the session for group G2, indicating that the repeated sessions helped the children from group G2 to look more at their partners. We do not observe an increase on the participants from group G1 (Figure 6.4a). However, the participants from group G1 seemed to look more at the partner in session H5 than in the other sessions, even if this result is not significant. This may indicate that the repeated sessions with the robot helped those children to look more at a human partner.

Smiles No effect of the groups was found on the smiling behavior of the children.



(a) Phase (a)



(b) Phase (b)

Figure 6.4: Observed behaviors by groups G1 and G2 through the sessions in Phases (a) and (b). H# for a session with a human partner, and R# for a session with Nao as partner, with # from 0 to 6 for each number of sessions.

Number of movements for the children to imitate We found a significant main effect of the groups on the number of the imitated movements ($F(1,9) = 34.94$; $p < 0.01$). Children from group G1 produce fewer imitation movements than the children from group G2 (G1: $M = 51.34$, $SD = 17.25$; G2: $M = 87.14$, $SD = 35.04$). This indicates that children from group G2 improved faster than children from group G1. Both groups had an increase in their number of movements.

Imitation scores We found a significant main effect of the scores on the groups ($F(1,9) = 17.65$; $p < 0.01$). Children from group G2 were more successful than children from group G1 (G1: $M = 23.09$, $SD = 26.81$; G2: $M = 79.79$, $SD = 20.25$). In Figure 6.4a, even if this result is not significant, we can observe that the children from G1 seemed to have a higher score in the sessions with the human partner, which is different from the children in group G2.

6.3.2 Phase (b)

In this phase, the partner had to imitate the arm movements of the child. Figure 6.4b shows the observed behaviors for the G1 and G2 groups in Phase (b). The participants from G1 had more difficulties understanding that they had to produce movements regardless of those of the partner, and that the robot was imitating them. When the session was with the robot, they often imitated the robot instead of initiating movements. Indeed, the robot was slightly moving from the noise based on skeleton data form the RGB-D camera and some children were reproducing the slight movements of the robot.

Gaze to the partner We observed a weak significant effect of the groups on the gaze direction of the children during Phase (b) ($F(1,5) = 5.46$; $0.05 < p < 0.1$). Participants from G1 looked globally less often to the robot than the participants from group G2 (G1: $M = 54.46$, $SD = 4.39$; G2: $M = 71.60$, $SD = 6.31$).

Smiles No effect of the groups was found on the smiling behavior of children. In Figure 6.4b children in group G1 smiled less in the last session of the therapy than in the other sessions.

Free initiation Free initiation was not significantly different between the groups on the whole experiment ($F(1,5) = 0.08$; $p = ns$). In Figure 6.4b, we can observe that in H5, the free initiation increased in both groups G1 and G2, and more drastically in G2 than in G1.

6.3.3 Description by Participants

We describe here the behaviors of the participants during the sessions. Figure 6.5 show the observed data for each participant for Phases (a) and (b).

CH3 (G1) CH3 was very enthusiastic to play with the robot and really appreciated the songs. In Phase (a), we observed an improvement of the gaze for the human partner after the sessions with the robot. He looked the human partner more successfully after the sessions with the robot. Indeed, in session H0, he gazed 80.52% of the time to the human, and 94.64% in session H5 after the four sessions with the robot partner and the week-long break. The number of movement to imitate increased with the sessions, indicating an improvement in the imitation. In Phase (b), the child was moving his arms to make the partner move, but he was not looking at the partner most of the time. Instead, he was closing his eyes or putting the head down.

CH5 (G1) This child had troubles understanding the tasks in Phases (a) and (b). In Phase (a), she gazed at her partner approximately 76% of the time. She imitated the human partner better than the robot partner. In Phase (b), it was very difficult to keep her focused on the task.

CH8 (G1) CH8 had difficulties understanding the imitation and initiation tasks in both Phases (a, b). However, his caregiver said he liked the sessions because he was not reluctant to come, and that he may have understood the task but was unwilling to perform the task.

CH10 (G1) This child performed the task but had some difficulties imitating the movement in Phase (a) (he was performing a random movement or always performing the same movement). At the end of each session with the robot, he kissed it on the head. The gaze behavior of this child did not improve with time in Phase (a); instead, it stayed stable. In Phase (b), the child

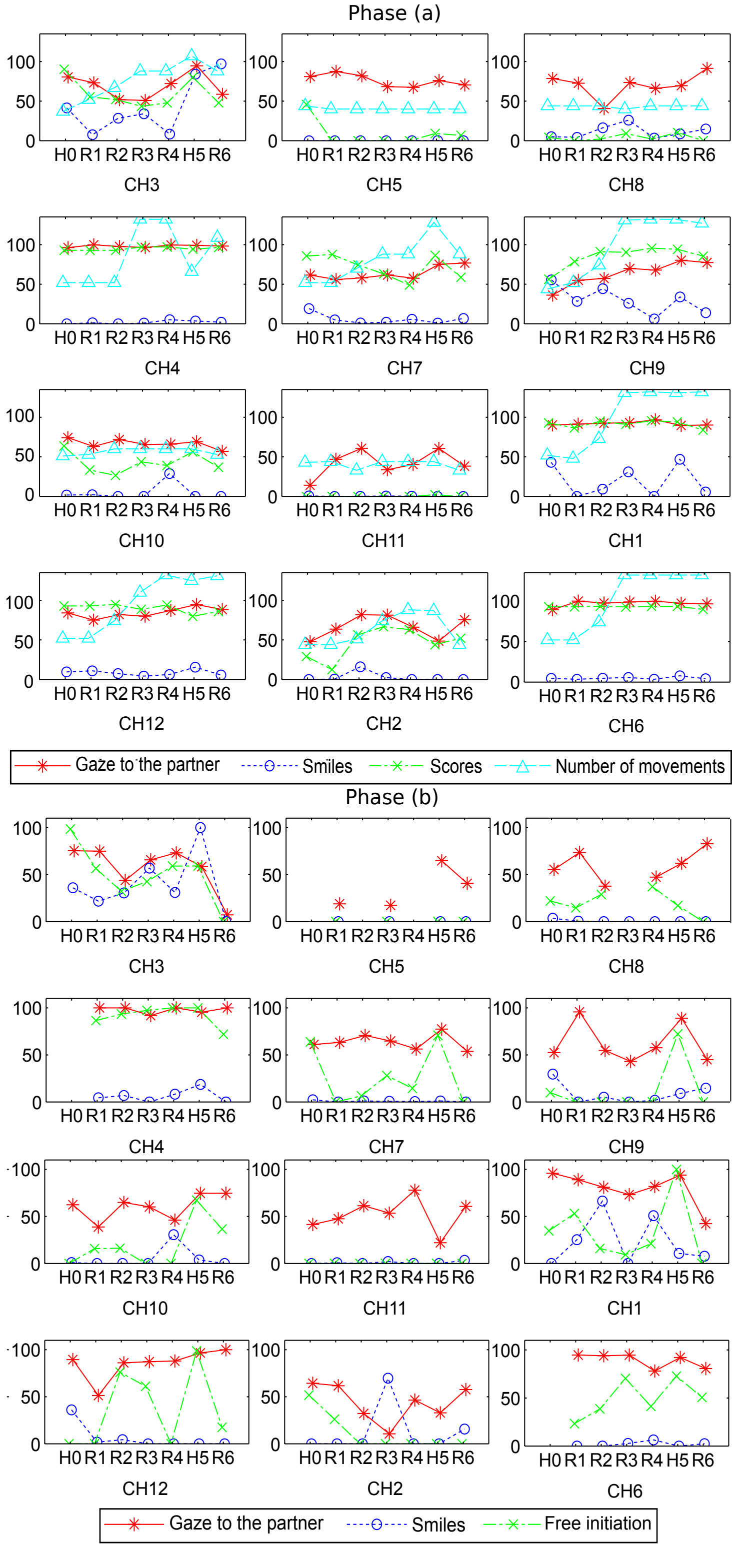


Figure 6.5: Observed behaviors of all the participants through the sessions in Phases (a) (top) and (b) (bottom). H# for a session with a human partner, and R# for a session with Nao as partner, with # from 0 to 6 for each number of sessions.

initiated movements for a significant amount of time in sessions H5 and R6. He may have benefited from the repetitive sessions with the robot for this skill.

CH11 (G1) CH11 had difficulties understanding the imitation and initiation tasks. He performed evading gaze behavior. However, his gaze at the partner was higher after session H0 in Phase (a). He gazed more at the human partner after the session with the robot. In H0, he looked at the partner 14.10% of the time, which was 60.72% in H5. By contrast, in Phase (b), his gaze increased from H0 to R4 and decreased with the human partner in H5.

CH1 (G2) CH1 was very enthusiastic about participating to the sessions. He even ran to the experiment room. In Phase (a), his gaze to the partner was good and constant over the time. He always succeeded with the imitation task and was able to follow the robot movement when the gestures were every 2 seconds. The number of movement to imitate increased with the sessions, indicating an improvement in the imitation. In Phase (b), his gaze to the robot decreased in the last session. Free initiation was higher in session H5 than in the other sessions.

CH4 (G2) The child expressed some reluctance to participate in the sessions and most of the time he asked to perform only two to three dances with the robot. However, his gaze on the partner was good and constant over the time. He always succeeded in the imitation task and he was able to follow the robot movement when the gestures were every 2 seconds. In Phase (b), the child did not want to initiate movement with the human partner in the first session H0, but he accepted the participation conditions during the remainder of the experiment. He initiated movement with the robot and tested the behavior of the robot. During session H5, he seemed to enjoy initiating movements with the human partner, and smiled more during this session.

CH7 (G2) This child was enthusiastic about performing the task at the beginning of a session and then showed signs of boredom from the middle to end of sessions. We observed that this participant increased his gaze on the partner in Phase (a). The number of movement to imitate increased with the sessions, indicating an improvement in the imitation. In Phase (b), with the robot partner, he did not understand he was imitated. He was imitating the robot movement. The percentage of free initiation was higher with the human partner ($M = 67.35$; $SD = 3.51$) than with the robot partner ($M = 9.81$; $SD = 11.80$), and no improvements were visible from H0 to H5 sessions. However, the child looked more at session H5 (70.86%) than at session H0 (63.85%).

CH9 (G2) In Phase (a), this child showed an improvement in her gazing behavior to the partner throughout the experiment. She also improved her imitation skills by being more focused on the partner. Indeed, she looked more at her caregiver than at the partner and was not focused on the task at all. We observed that this participant increased his gaze on the partner in Phase (a). The number of movement to imitate increased with the sessions, indicating an improvement in the imitation. She succeeded in reaching the level at which were the partner was moving every 2 seconds. In Phase (b), she did not initiate the partner except in session H5. She may have benefited from the repetitive sessions with the robot for this skill.

CH12 (G2) This child was enthusiastic to participate in the sessions. He understood the tasks. In Phase (a), we observed that his gaze towards the partner increased through the sessions. The gaze on the human after the sessions with Nao robot was higher compared to during the first session. The number of movement to imitate increased with the sessions, indicating an improvement in the imitation. He succeeded in performing the imitation task and was able to follow the robot movement when the gestures occurred every 2 seconds. In Phase (b), except in session R1, the participant focused his attention on the partner. The gaze percentage increased over time. In session H5, he was really enthusiastic to initiate movement in the human partner and asked to repeat the task several time.

CH2 (G3) As mentioned before, this child was very agitated and perturbed for the last four sessions due to personal issues. From H0 to R2, the gaze to the partner of the child increased and then decreased for Phase (a). In Phase (b), he was quite agitated and it was difficult to get the robot to imitate him.

CH6 (G3) CH6 was very enthusiastic about participating in the sessions. He even ran to the experiment room and asked if he was going to see the robot when he saw the experimenter in the care center. In Phases (a) and (b), his gaze on the partner was good and constant over time. He always succeeded in performing the imitation task and was able to follow the robot movement when the gestures were every 2 seconds. The number of movement to imitate increased with the sessions, indicating an improvement in the imitation. In Phase (b), his gaze on the robot decreased in the last session. He had few difficulties understanding Phase (b) with the robot partner. He imitated the robot in the beginning, but after a few trials, he understood the task and initiated movement for the robot.

6.4 Conclusion

With the experiment that we described in this Chapter, we aimed to develop a robot-based imitation experimental protocol for children with ND. We aimed to validate the relevance of the proprioceptive and visual individual profiles with the hypothesis that hyporeactivity to visual motion and an overreliance on proprioceptive information would participate to difficulties in integrating interaction cues and engaging in successful interactions. We designed an imitation task for seven repeated sessions across a period of eight weeks. In the first and sixth session, the child's partner was a human; in the other sessions the partner was a robot, Nao. There were two phases in a session: Phase (a) the child imitated the partner and Phase (b) the partner imitated the child.

We observed that participants who had hyporeactivity to visual motion and an overreliance on proprioceptive information (children from group G1) gazed less towards the partner than participants relying on visual information (children from group G2), during both phases of the sessions. Children from group G1 were less successful in imitating the partner than children from G2. This validates our hypothesis **H1A**. Children from group G1 were less likely to initiate movement that could be imitated by the partner than children from group G2, which validates **H1B**. Finally, children from group G1 had a slower improvement in their imitation skills than children from group G2, which validates our hypothesis **H1C**. We observed that children from group G2 gazed more towards their partners during the sessions in Phase (a). Children from group G1 seemed to gaze more towards the human partner in the sixth session than in the other sessions in Phase (a). The repeated interactions may have helped the children in both groups to look more at the human partner. In Phase (b), we observed that children from group G2 initiated more movements for the human partner in the sixth session than in the other sessions in Phase (b). This may indicate that the repetitive sessions with the robot improved their will to initiate movements with a human partner. However, some children did not succeed in imitating the partner across the sessions or did not seem to understand that they were being imitated. The task was maybe too complicated or inadequately adapted, and they might have needed more sessions to observe improvements in their imitation skills. For some children, Phase (b) was difficult to set-up, because they were unwilling to stand in front of the RGB-D camera for the time when they were correctly tracked, which might have contributed to their limited understanding of this phase. Overall, the repetitive sessions seemed to have a positive impact on the children' interaction skills belonging in both groups G1 and G2. In addition, children seemed to enjoy the interactions: most of them showed enthusiasm for interacting with the robot.

In [Srinivasan et al., 2015], the authors found that repeated session with a robot to train imitation helped children with ASD to improve their imitation and motor skills. As them, we found encouraging results of the use of robot to improve imitation in children with ND. As in [Tapus et al., 2012], we observed inter-differences between the children with ASD: some children had more difficulties in imitating or to engage with the robot than others. In our work, we observed that these difficulties were linked to their proprioceptive and visual profile. This can enlighten the need of profiles prior to propose a robot-based therapy to children with ND.

In [Haswell et al., 2009] and [Izawa et al., 2012], the authors observed that the greater the reliance on proprioception, the greater the child's impairments in social function and imitation. We observed consistent results in this experimental setup. In addition, in [Salowitz et al., 2013] the authors observed that in children with ASD, visuospatial information processing deficits may contribute importantly to functional motor coordination impairment. They observed that children with ASD were less accurate than TD individuals in imitation of hand shape and orientation, and the number of consistent limb movements. In our study, we also observed that children with ND showing hyporeactivity to visual cues showed more imitation impairments.

In this chapter, we observed that the proprioceptive and visual profiles of the participants predicted their behaviors. We also observed that the repeated sessions with the robot increased the interaction behaviors of the children towards a human partner.

Chapter 7

Conclusions and Perspectives

7.1 Reminder of the PhD Thesis Objectives

The long term goal of this PhD Thesis is the development of new models of personalized robot-based social interactions for individuals with ND. Robot-based therapy has been observed to be efficient to enhance social skills in individuals with ASD [Hart, 2005]. However, there are strong inter-individual differences among individuals with ASD, and finding an appropriate way to define an individualized model is a major concern. As we mentionned in the introduction, our participants are patients with more general Neurodevelopmental disorders, so we based our work on inter-individual differences with respect to the sensory integration by individuals with ND. To define relevant individual profiles, we used the hypothesis that an individual's reliance on proprioceptive and kinematic visual cues would affect the way he/she interacts with a humanoid robot [Haswell et al., 2009, Liu and Chepyator-Thomson, 2009, Saracho, 2003, Coates et al., 1975]. The hypothesis driving our work was that:

- **H1** A mitigated behavioral response (i.e., hyporeactivity) to visual motion and an overreliance on proprioceptive information are linked in individuals with ND to their difficulties in integrating social cues and engaging in successful interactions.

We worked in collaboration of three care-center, and worked with a subject-pool composed of 19 participants: children, teenagers, and adults with ND.

7.2 Summary of the PhD Thesis

The contributions of the PhD thesis are:

The experiments were done with the same participants over a period of two years

A method to define proprioceptive and visual profiles of individuals with ND

In chapter 2, we proposed a method to define the profiles of our participants. Our method combines a well-known questionnaire, the AASP [Brown and Dunn, 2002] (see Appendix A), and an experimental setup. Thanks to this method, we were able to identify the sensory integration preferences of each of our participants, in regards of visual and proprioceptive cues. We observed that in individuals with ND, the age did not drive the postural response to a visual stimulus, in contrast of TD individuals. In our subject-pool, we extracted three groups of different visual and proprioceptive preferences.

- Group G1 in which the participants showed an overreliance on proprioceptive cues and hyporeactivity to visual cues.
- Group G2 in which the participants showed an even reliance on visual and proprioceptive cues.
- Group G3 in which the participants showed a hyporeactivity to proprioceptive cues and an overreliance on visual cues.

We expected that participants in the most proprioceptive group (i.e., G1) would display more difficulties to interact and to enhance their communication skills. On the contrary, the most visual group (i.e., G3) would interact more easily with the robot and would enhance their communication skills.

The validation of the proprioceptive and visual profiles as an indicator of the social behavior of individuals with ND

We tested if the three groups enabled us to predict the social and communication behavior of our participants, in four social and communication skills (i.e., emotion recognition, short greeting task, answer to joint attention, and imitation). We also observed if the groups enabled us to predict the learning faculties of our participants in an imitation task with the robot over repeated sessions across eight weeks.

In emotion recognition

We started assessing the potential link between recognition performances of body/facial expressions of emotions on embodiment with different visual features (two humanoid robots, a virtual agent, and a human), and proprioceptive and visual cues integration profiles of our participants (see Chapter 3). We observed if the sensory preferences had an impact on emotion recognition, but also if they induced a preference to a particular channel of expression (body/facial/both) or embodiment (robots/virtual agent/human). First, we collected a database (EMBODI-EMO) and validated it with TD individuals. Then, we used it to investigate the relationship between emotion recognition and proprioceptive and visual profiles of our 19 participants with ND. We found that adults with ND had more difficulties to perform the task than children with ND. We observed that the integration of visual and proprioceptive cues of a children with ND influences his/her abilities in recognizing emotions. Children with ND relying on proprioceptive information had lower recognition scores than the other children with ND. We did not observe a favorite channel of expression of emotion, neither a favorite embodiment. The task was difficult for our participants with ND. Indeed, we worked with participants with various level of impairments, and some had more difficulties to perform the task. After this experiment, we searched to develop an easy more interactive task with the robot, that could be understood and enjoyed by all our participants.

In a short greeting task with a robot

In chapter 4, we introduced our participants with ND to the Nao humanoid robot. The main objective of this first interaction was to present the robot as a social partner and to remove the novelty effect of the robot for further experiment. We also wanted the participants to be more at ease with the robot, as individual with ND can show stress or reluctance to new things. We designed it as a short greeting task, in which the robot introduced itself to the participant as his/her new friend, and performed a dance for him/her. In addition, we observed the social behavior of our participants towards the robot, and observed if there was a link with their profiles. Overall, most of the participants showed a strong interest towards the robot. Some participants were slightly afraid of and impressed by the robot, but these participants seemed to be reassured at the end of the interaction. We observed that participants from group G3 showed more free speech than participants from groups G1 and G2, and participants from group G1 were the less talkative. Furthermore, there was no statistical evidence that the most visual participants looked more at the robot during the interaction than the other participants, yet we did observe a tendency. This interaction was a encouraging start to validate our hypothesis.

In a joint attention task with a robot

With the experiment that we described in chapter 5, we aimed to assess the response of our participants to a joint attention initiation. To do so, we designed a matching game with the Nao robot. We observed that participants with an overreliance on proprioceptive cues and hyporeactivity to visual cues missed more the prompting and seemed to follow the prompting of the robot more slowly than individuals with an overreliance on visual cues and a hyporeactivity to proprioceptive cues. However, even if we talked and designed the interaction with the caregivers, the matching game was difficult for participants with low cognition. We observed signs of boredom at the end of the sessions for some participants.

In an imitation task with the robot over eight weeks period

As a follow-up to our previous experiments, we were confident in the fact that the profiles based on the integration of proprioceptive and visual integration of cues might be an accurate predictor of the social skills of our participants with ND. We continued our approach research with an imitation task with the robot over eight weeks.

In chapter 6, we described the design of an experimental protocol to assess and enhance imitation skills of the children with ND. The experimental protocol was composed of seven imitation sessions over eight weeks. We aimed to assess the capabilities of our children participants to imitate or to initiate gestures of a human or robot partner. During the first and sixth session, the partner was human, and on the rest of the sessions the Nao robot. As expected, we observed that children with an overreliance on proprioceptive cues and hyporeactivity to visual cues had more difficulty imitating the robot than the other children. Moreover, most of the children exhibited some positive effects in their social behavior (gaze towards the partner, imitations) toward a human partner after the sessions with the robot. However, children with an overreliance on proprioceptive cues and hyporeactivity to visual cues showed difficulties to perform the task. In addition, children seemed to enjoy the interactions: most of them showed enthusiasm for interacting with the robot, even if they had difficulties to accomplish the task.

Conclusion

In conclusion, the profiles we defined appear to be a good predictor of the social and communication skills and the learning abilities of our participants with ND. However, participants with the most low-cognition level had difficulties to interact with the robot and understand the tasks. We also observed that the interaction with a robot was pleasant for individual with ND, and that skills learned with the robot seemed to be generalized to a human partner. Another limitation of our work is that we did not have access to the diagnostic of our participants. Refering to them as having neurodevelopemental disorder should be more exact.

7.3 Perspectives

Future works should focus on the design of an autonomous robot-based model, personalized to the users' sensory preferences and performance. Indeed, a main interest in SAR is to propose more autonomous robots, that will be able to adapt to the individual needs of children over longer periods of time [Thill et al., 2012].

Numerous studies on HRI and SAR are working on the development of more autonomous methods to enable more autonomous and adaptive interactions for the users. The development of tools tracking the behavior of the user can provide helpful information for the robot behavior during the interaction. The data can provide information on the engagement status (from RGB-D camera [Anzalone et al., 2015] or multimodal features [Vaufreydaz et al., 2016]) or the emotional status [Barros et al., 2015, Patwardhan and Knapp, 2016] of the user, which would enable the robot to adopt the correct behavior towards the user. Tracking the behavior can also be part of the task to perform with the robot. For example, in [Tapus et al., 2012, Ranatunga et al., 2013, Lei et al., 2015], RGB-D cameras, as the Kinect or Xtion, were used to track part of full-body motion, for imitation tasks for children with ASD. Force sensors were used in [Dautenhahn et al., 2009a, Costa et al., 2015] to track and correct the touch sense in children with ASD.

In SAR for individual with ASD, other sensory preferences can also be taken in account in the development of robot-based therapy. Individuals with ASD have great inter-individual variability in their symptoms and in their sensory integration of cues [APA, 2013]. As saw in this PhD thesis report, individuals with ASD showed great variability in vision, motor, and proprioception [Gowen and Hamilton, 2013, Simmons et al., 2009]. Moreover, individuals with ASD show an unusual integration of audio information [Rogers and Ozonoff, 2005] and touch [Kaiser et al., 2016]. In [Čeponienė et al., 2003], the authors observed that individuals with ASD have been shown to perceive well music and to outperform their peers in pitch discrimination and the perception of the detailed structure of melodies segments, but have impairments in processing speech. Touch is also an important sense in the learning of social and affective behavior [Voos et al., 2013], and is impaired in individuals with ASD [Kaiser et al., 2016].

The sensory preferences of individual with ASD is more and more taken into account [Simmons et al., 2009, Hilton et al., 2010, Gowen and Hamilton, 2013], and have been outlined in SAR for ASD [Ferrari et al., 2009]. Indeed, in SAR, the robot should not be overwhelming by providing too many sensory inputs for the individuals with ASD [Ferrari et al., 2009].

The use of the sensory preferences in individual with ASD in a robot-based therapy could be used in different examples (e.g., to learn social movement). Individuals with ASD have difficulties to follow and understand social and biological movement as indicated in [Simmons et al., 2009]. A robot showing more and more social movement across the session and improvement of the user could be appropriate to individual showing an hyporeactivity to visual cues. Individuals with ASD could also work on the touch and body perception with a robot, as already done in [Dautenhahn et al., 2009a, Costa et al., 2015]. The authors use KASPAR robot to learn the children with ASD to control their touch in order to not hurt the robot and to identify body parts.

In addition, robots are called to be more and more present in our daily life and in SAR. The adaptation of the robot behavior on the user's personality is an important topic of interest [Tapus et al., 2008, Castellano et al., 2009]. The use of sensory preferences in individuals (TD or not) could be an interesting research topic. As saw in the chapter 1, the dependence to the visual field leads the social behavior in TD individuals [Liu and Chepyator-Thomson, 2009, Saracho, 2003]. The authors of [Robertson and Simmons, 2013] tested the autistic traits, thank to the Autism-Spectrum Quotient [Baron-Cohen et al., 2001b], and the sensory preferences, thanks to their own questionnaire, of a TD population. They observed a strong link between sensory processing and autistic traits. These results could enlighten the implication of sensory processing deficiencies in social interaction difficulties.

A future work would be to implement a framework for mulimodal generation [Kopp et al., 2006, Scherer et al., 2012] for a more autonomous robot which will detect and adapt to the proprioceptive and visual profiles. We propose the framework shown in Figure 7.1. The tracking of the of the participant's behavior would enable us to compute his/her profile. It would impact the robot behavior (e.g.: moving slowly; moving the face and/or the body). The task will also adapt to the profile and the performance of the participants (e.g.: the task would be easier and the difficulty will increase more slowly if the participant is relying heavily on proprioceptive cues). Visual information from a RGB-D camera would give information on the gaze behavior and the engagement of the participant, and would give knowledge on the perfomance. This would enable to adapt the task level and the behavior of the robot (e.g.: encouraging the participant, asking for attention...).

In the JA task described in chapter 5, the framework we designed was not taking in account the integration of the proprioceptive and visual profile, but the visual informations were used to lead the task. Unfortunatly, due to difficulty to track the participant, the experimenter was

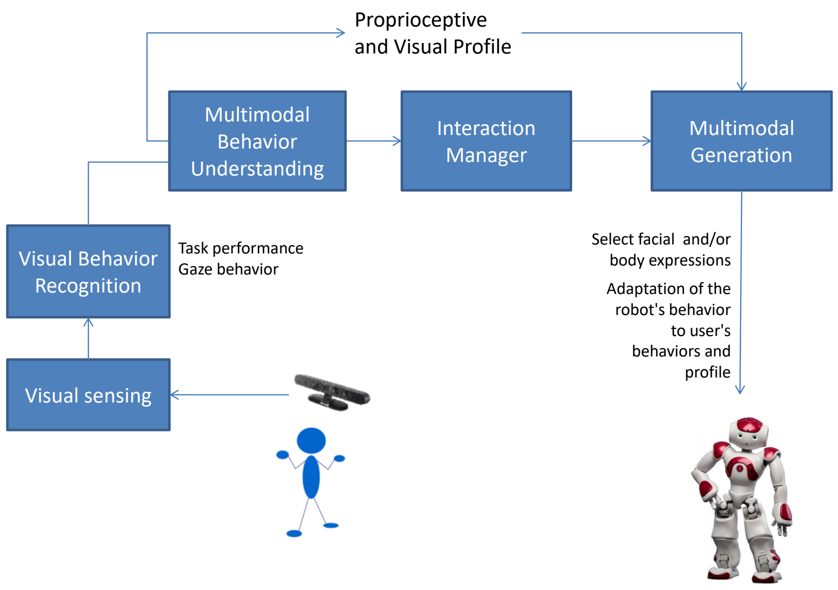


Figure 7.1: Proposed Framework. With this framework, the RGB-D data would enable the system to detect the profiles, the performance and the social behavior of the participants. With these, the system would modify the behavior of the robot accordingly to the participant's information.

overriding the framework. In the imitation task described in chapter 6, the Phase (a) (i.e.: the participant imitate the robot) was leaded by the experimenter. The next step would be to develop a movement recognition module and a performance tracker to lead the interaction. In the Phase (b), the robot was tracking the movement of the participant. Tracking the engagement would enable it to end the interaction or to ask for attention when needed.

7.4 General Conclusion

To conclude this PhD Thesis, using individual profiles to adapt robot-based therapies seems to be an adequate method for individuals with ND. We observed again that robot-based therapy was pleasant for most of our participants with ND.

We were able to observe that the repeated sessions with the robot Nao were beneficial for participants with ND: after the sessions with the robot Nao, the participants showed an improvement in their social skills (gaze to the partner, imitations) towards a human partner.

Defining such individual profiles could provide promising strategies for designing successful and adapted Human-Robot Interaction for individuals with ND.

A limitation of our work was the lack of access to diagnostic reviews as ADOS-G, ADI-R or other ASD diagnostic assessments of our participants.

Appendix A

Adolescent and Adult Sensory Profile (AASP)

The AASP developed by Dunn [Brown and Dunn, 2002] is a questionnaire that enable to collect an individual's movement, visual, touch, auditory, taste, and activity level sensory processing preferences. The sensory processing preferences are described in terms of the quadrants in Dunn's model of sensory processing [Dunn, 1999]: low registration, sensation seeking, sensory sensitivity, and sensation avoiding.

These quadrants are formed by the junctions between individual differences in neurological thresholds for stimulation (high-low) and self-regulation strategies (active-passive):

- Low registration (high-passive): tendency to miss or take a long time to respond to stimuli.
- Sensation Seeking (high-active): tendency to try to create additional stimuli or to look for environments that provide sensory stimuli.
- Sensory Sensitivity (low-passive): tendency to answer quickly to stimuli.
- Sensation Avoiding (low-active): tendency to be overwhelmed or bothered by sensory stimuli and to be actively involved to reduce the stimuli in the environment.

Table A.1: Examples of the Sensory Profiles items

I enjoy going to environment were the lights and colors are bright
I am bothered when I see lots of movements around me
I jump with surprise easily when I hear sudden or intense noise

As it is a commercial questionnaire, we cannot provide the list of the 29 items that were filled by our participants. We provide some examples in Table A.1. For each item, the person filling the form has to rate from 1 to 5 if the behavior described occurs to him/her with '1' is never and '5' is all the time.

Appendix B

Robotic and Virtual Agent Platforms Used

B.1 Nao Robot

The humanoid NAO robot (Figure B.1) is developed by SoftBank Robotics[1], former Aldebaran Robotics. This robot has 25 degrees of freedom and is equipped with many sensors that allow it to perceive its surrounding environment with high precision and stability. Its face is static.

This robot is employed in the studies of Chapters 3, 4, 5 and 6.

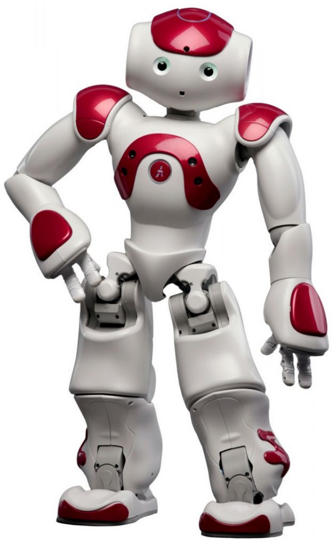


Figure B.1: Humanoid Nao robot

[1] https://www.ald.softbankrobotics.com/fr

B.2 Zeno Robot

The humanoid R25 (Zeno) robot (Figure B.2) is developed by Robokind[2]. This robot has a full-motion body and an expressive face. The face of the robot composed of synthetic skin, which enables it to express emotions on its face. Zeno has five degrees of freedom in his face, with which he can make basic facial expressions. In addition, the robot can shift his gaze using two degrees of freedom.

This robot is employed in Chapter 3.

Figure B.2: Humanoid Zeno robot

[2] http://www.robokindrobots.com/support-documentation/r25/

B.3 Mary Virtual Agent

The virtual agent Mary (Figure B.3) comes from the Multimodal Affective and Reactive Character (MARC) platform[3] [Courgeon and Clavel, 2013], and is a female virtual agent. She can display several postures, facial expressions and behaviors.

This virtual agent is used in Chapter 3.

Figure B.3: Mary virtual agent

[3] http://www.marc-toolkit.net/?page=home

Bibliography

[Anzalone et al., 2015] Anzalone, S. M., Boucenna, S., Ivaldi, S., and Chetouani, M. (2015). Evaluating the engagement with social robots. *International Journal of Social Robotics*, 7(4):465–478.

[Anzalone et al., 2014] Anzalone, S. M., Tilmont, E., Boucenna, S., Xavier, J., Jouen, A.-L., Bodeau, N., Maharatna, K., Chetouani, M., Cohen, D., Group, M. S., et al. (2014). 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*, 8(7):814–826.

[APA, 2013] APA (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

[Ashwood et al., 2015] Ashwood, K. L., Buitelaar, J., Murphy, D., Spooren, W., and Charman, T. (2015). European clinical network: autism spectrum disorder assessments and patient characterisation. *European child & adolescent psychiatry*, 24(8):985–995.

[Assaiante, 1998] Assaiante, C. (1998). Development of locomotor balance control in healthy children. *Neuroscience & biobehavioral reviews*, 22(4):527–532.

[Atkinson, 2009] Atkinson, A. P. (2009). Impaired recognition of emotions from body movements is associated with elevated motion coherence thresholds in autism spectrum disorders. *Neuropsychologia*, 47(13):3023–3029.

[Bagust et al., 2013] Bagust, J., Docherty, S., Haynes, W., Telford, R., and Isableu, B. (2013). Changes in rod and frame test scores recorded in schoolchildren during development–a longitudinal study. *PloS one*, 8(5):e65321.

[Baio, 2012] Baio, J. (2012). Prevalence of autism spectrum disorders: Autism and developmental disabilities monitoring network, 14 sites, united states, 2008. morbidity and mortality weekly report. surveillance summaries. volume 61, number 3. *Centers for Disease Control and Prevention*.

[Baird et al., 2006] Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., and Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in south thames: the special needs and autism project (snap). *The lancet*, 368(9531):210–215.

[Bar-Haim and Bart, 2006] Bar-Haim, Y. and Bart, O. (2006). Motor function and social participation in kindergarten children. *Social Development*, 15(2):296–310.

[Baranek et al., 2013] Baranek, G. T., Watson, L. R., Boyd, B. A., Poe, M. D., David, F. J., and McGuire, L. (2013). Hyporesponsiveness to social and nonsocial sensory stimuli in children with autism, children with developmental delays, and typically developing children. *Development and Psychopathology*, 25(02):307–320.

[Baron-Cohen et al., 2009] Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., and Brayne, C. (2009). Prevalence of autism-spectrum conditions: Uk school-based population study. *The British Journal of Psychiatry*, 194(6):500–509.

[Baron-Cohen et al., 2000] Baron-Cohen, S., Wheelwright, S., Cox, A., Baird, G., Charman, T., Swettenham, J., Drew, A., and Doehring, P. (2000). Early identification of autism by the checklist for autism in toddlers (chat). *Journal of the royal society of medicine*, 93(10):521.

[Baron-Cohen et al., 2001a] Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001a). The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of child psychology and psychiatry*, 42(2):241–251.

[Baron-Cohen et al., 1997] Baron-Cohen, S., Wheelwright, S., and Jolliffe, T. (1997). Is there a" language of the eyes"? evidence from normal adults, and adults with autism or asperger syndrome. *Visual Cognition*, 4(3):311–331.

[Baron-Cohen et al., 2001b] Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., and Clubley, E. (2001b). The autism-spectrum quotient (aq): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of autism and developmental disorders*, 31(1):5–17.

[Barros et al., 2015] Barros, P., Jirak, D., Weber, C., and Wermter, S. (2015). Multimodal emotional state recognition using sequence-dependent deep hierarchical features. *Neural Networks*, 72:140–151.

[Begeer et al., 2006] Begeer, S., Rieffe, C., Terwogt, M. M., and Stockmann, L. (2006). Attention to facial emotion expressions in children with autism. *Autism*, 10(1):37–51.

[Begum et al., 2015] Begum, M., Serna, R. W., Kontak, D., Allspaw, J., Kuczynski, J., Yanco, H. A., and Suarez, J. (2015). Measuring the efficacy of robots in autism therapy: How informative are standard hri metrics'. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, pages 335–342. ACM.

[Bekele et al., 2014] Bekele, E., Crittendon, J. A., Swanson, A., Sarkar, N., and Warren, Z. E. (2014). Pilot clinical application of an adaptive robotic system for young children with autism. *Autism*, 18(5):598–608.

[Billard, 2003] Billard, A. (2003). Robota: Clever toy and educational tool. *Robotics and Autonomous Systems*, 42(3):259–269.

[Billard et al., 2007] Billard, A., Robins, B., Nadel, J., and Dautenhahn, K. (2007). Building robota, a mini-humanoid robot for the rehabilitation of children with autism. *Assistive Technology*, 19(1):37–49.

[Bondy and Frost, 1994] Bondy, A. S. and Frost, L. A. (1994). The picture exchange communication system. *Focus on Autism and Other Developmental Disabilities*, 9(3):1–19.

[Boraston and Blakemore, 2007] Boraston, Z. and Blakemore, S.-J. (2007). The application of eye-tracking technology in the study of autism. *The Journal of physiology*, 581(3):893–898.

[Bray et al., 2004] Bray, A., Subanandan, A., Isableu, B., Ohlmann, T., Golding, J. F., and Gresty, M. A. (2004). We are most aware of our place in the world when about to fall. *Current Biology*, 14(15):R609–R610.

[Brown and Dunn, 2002] Brown, C. and Dunn, W. (2002). *Adolescent-Adult Sensory Profile: User's Manual*. Therapy Skill Builders.

[Buisine et al., 2014] Buisine, S., Courgeon, M., Charles, A., Clavel, C., Martin, J.-C., Tan, N., and Grynszpan, O. (2014). The role of body postures in the recognition of emotions in contextually rich scenarios. *International Journal of Human-Computer Interaction*, 30(1):52–62.

[Burke et al., 1976] Burke, D., Hagbarth, K.-E., Löfstedt, L., and Wallin, B. G. (1976). The responses of human muscle spindle endings to vibration of non-contracting muscles. *The Journal of physiology*, 261(3):673–693.

[Carpenter et al., 2005] Carpenter, M., Tomasello, M., and Striano, T. (2005). Role reversal imitation and language in typically developing infants and children with autism. *Infancy*, 8(3):253–278.

[Cascio et al., 2012] Cascio, C. J., Foss-Feig, J. H., Burnette, C. P., Heacock, J. L., and Cosby, A. A. (2012). The rubber hand illusion in children with autism spectrum disorders: delayed influence of combined tactile and visual input on proprioception. *Autism*, 16(4):406–419.

[Castellano et al., 2009] Castellano, G., Pereira, A., Leite, I., Paiva, A., and McOwan, P. W. (2009). Detecting user engagement with a robot companion using task and social interaction-based features. In *Proceedings of the 2009 international conference on Multimodal interfaces*, pages 119–126. ACM.

[Celani et al., 1999] Celani, G., Battacchi, M. W., and Arcidiacono, L. (1999). The understanding of the emotional meaning of facial expressions in people with autism. *Journal of autism and developmental disorders*, 29(1):57–66.

[Čeponienė et al., 2003] Čeponienė, R., Lepistö, T., Shestakova, A., Vanhala, R., Alku, P., Näätänen, R., and Yaguchi, K. (2003). Speech–sound-selective auditory impairment in children with autism: they can perceive but do not attend. *Proceedings of the National Academy of Sciences*, 100(9):5567–5572.

[Chamak et al., 2011] Chamak, B., Bonniau, B., Oudaya, L., and Ehrenberg, A. (2011). The autism diagnostic experiences of french parents. *Autism*, 15(1):83–97.

[Chang et al., 2014] Chang, Y.-S., Owen, J. P., Desai, S. S., Hill, S. S., Arnett, A. B., Harris, J., Marco, E. J., and Mukherjee, P. (2014). Autism and sensory processing disorders: shared white matter disruption in sensory pathways but divergent connectivity in social-emotional pathways. *PloS one*, 9(7):e103038.

[Charman and Baron-Cohen, 1994] Charman, T. and Baron-Cohen, S. (1994). Another look at imitation in autism. *Development and psychopathology*, 6(03):403–413.

[Charman et al., 1997] Charman, T., Swettenham, J., Baron-Cohen, S., Cox, A., Baird, G., and Drew, A. (1997). Infants with autism: an investigation of empathy, pretend play, joint attention, and imitation. *Developmental psychology*, 33(5):781.

[Chiari et al., 2002] Chiari, L., Rocchi, L., and Cappello, A. (2002). Stabilometric parameters are affected by anthropometry and foot placement. *Clinical Biomechanics*, 17(9):666–677.

[Clavel et al., 2009] Clavel, C., Plessier, J., Martin, J.-C., Ach, L., and Morel, B. (2009). Combining facial and postural expressions of emotions in a virtual character. In *International Workshop on Intelligent Virtual Agents*, pages 287–300. Springer.

[Coates et al., 1975] Coates, S., Lord, M., and Jakabovics, E. (1975). Field dependence-independence, social-non-social play and sex differences in preschool children. *Perceptual and Motor Skills*, 40(1):195–202.

[Constantino and Gruber, 2007] Constantino, J. N. and Gruber, C. P. (2007). *Social responsiveness scale (SRS)*. Western Psychological Services Los Angeles, CA.

[Conti et al., 2015] Conti, D., Di Nuovo, S., Buono, S., Trubia, G., and Di Nuovo, A. (2015). Use of robotics to stimulate imitation in children with autism spectrum disorder: A pilot study in a clinical setting. In *Robot and Human Interactive Communication (RO-MAN), 2015 24th IEEE International Symposium on*, pages 1–6. IEEE.

[Cooney et al., 2015] Cooney, S. M., O'Shea, A., and Brady, N. (2015). Point me in the right direction: Same and cross category visual aftereffects to directional cues. *PloS one*, 10(10):e0141411.

[Costa et al., 2015] Costa, S., Lehmann, H., Dautenhahn, K., Robins, B., and Soares, F. (2015). Using a humanoid robot to elicit body awareness and appropriate physical interaction in children with autism. *International journal of social robotics*, 7(2):265–278.

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

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

[Coulson, 2004] Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of nonverbal behavior*, 28(2):117–139.

[Courgeon and Clavel, 2013] Courgeon, M. and Clavel, C. (2013). Marc: a framework that features emotion models for facial animation during human–computer interaction. *Journal on Multimodal User Interfaces*, 7(4):311–319.

[Courgeon et al., 2012] Courgeon, M., Grynszpan, O., Buisine, S., and Martin, J. (2012). Interactive expressive virtual characters: challenges for conducting experimental studies about multimodal social interaction. In *Proceedings of the 9th International Conference on Disability, Virtual Reality & Associated Technologies Laval*.

[Courgeon et al., 2014] Courgeon, M., Rautureau, G., Martin, J.-C., and Grynszpan, O. (2014). Joint attention simulation using eye-tracking and virtual humans. *IEEE Transactions on Affective Computing*, 5(3):238–250.

[Crane et al., 2009] Crane, L., Goddard, L., and Pring, L. (2009). Sensory processing in adults with autism spectrum disorders. *Autism*, 13(3):215–228.

[Cunningham et al., 2006] Cunningham, D. W., Nusseck, H.-G., Teufel, H., Wallraven, C., and Bulthoff, H. H. (2006). A psychophysical examination of swinging rooms, cylindrical virtual reality setups, and characteristic trajectories. In *IEEE Virtual Reality Conference (VR 2006)*, pages 111–118. IEEE.

[Dael et al., 2012] Dael, N., Mortillaro, M., and Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion*, 12(5):1085.

[Dautenhahn et al., 2009a] Dautenhahn, K., Nehaniv, C. L., Walters, M. L., Robins, B., Kose-Bagci, H., Mirza, N. A., and Blow, M. (2009a). Kaspar–a minimally expressive humanoid robot for human–robot interaction research. *Applied Bionics and Biomechanics*, 6(3-4):369–397.

[Dautenhahn et al., 2009b] Dautenhahn, K., Nehaniv, C. L., Walters, M. L., Robins, B., Kose-Bagci, H., Mirza, N. A., and Blow, M. (2009b). Kaspar–a minimally expressive humanoid robot for human–robot interaction research. *Applied Bionics and Biomechanics*, 6(3-4):369–397.

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

[Dawson et al., 1990] Dawson, G., Hill, D., Spencer, A., Galpert, L., and Watson, L. (1990). Affective exchanges between young autistic children and their mothers. *Journal of abnormal child psychology*, 18(3):335–345.

[Dawson et al., 1998] Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., and Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of autism and developmental disorders*, 28(6):479–485.

[Dawson et al., 2010] Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., Donaldson, A., and Varley, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: the early start denver model. *Pediatrics*, 125(1):e17–e23.

[Dawson et al., 2004] Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., and Liaw, J. (2004). Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Developmental psychology*, 40(2):271.

[De Gelder et al., 2015] De Gelder, B., de Borst, A., and Watson, R. (2015). The perception of emotion in body expressions. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(2):149–158.

[De Gelder and Van den Stock, 2011] De Gelder, B. and Van den Stock, J. (2011). The bodily expressive action stimulus test (beast). construction and validation of a stimulus basis for measuring perception of whole body expression of emotions. *Frontiers in psychology*, 2:181.

[Diehl et al., 2012] Diehl, J. J., Schmitt, L. M., Villano, M., and Crowell, C. R. (2012). The clinical use of robots for individuals with autism spectrum disorders: A critical review. *Research in autism spectrum disorders*, 6(1):249–262.

[Dulac et al., 1994] Dulac, O., Chugani, H. T., and Dalla Bernardina, B. (1994). *Infantile spasms and West syndrome*. WB Saunders Company.

[Dunn, 1999] Dunn, W. (1999). The sensory profile: User's manual (psychological corporation, san antonio, tx).

[Duquette et al., 2008] Duquette, A., Michaud, F., and Mercier, H. (2008). Exploring the use of a mobile robot as an imitation agent with children with low-functioning autism. *Autonomous Robots*, 24(2):147–157.

[Dziuk et al., 2007] Dziuk, M., Larson, J., Apostu, A., Mahone, E., Denckla, M., and Mostofsky, S. (2007). Dyspraxia in autism: association with motor, social, and communicative deficits. *Developmental Medicine & Child Neurology*, 49(10):734–739.

[Edin, 2001] Edin, B. B. (2001). Cutaneous afferents provide information about knee joint movements in humans. *The Journal of Physiology*, 531(1):289–297.

[Ekman and Friesen, 1984] Ekman, P. and Friesen, W. (1984). Emotion facial action coding system (em-facs). *San Francisco: University of California*.

[Elsabbagh et al., 2012] Elsabbagh, M., Divan, G., Koh, Y.-J., Kim, Y. S., Kauchali, S., Marcín, C., Montiel-Nava, C., Patel, V., Paula, C. S., Wang, C., et al. (2012). Global prevalence of autism and other pervasive developmental disorders. *Autism Research*, 5(3):160–179.

[Erden, 2013] Erden, M. S. (2013). Emotional postures for the humanoid-robot nao. *International Journal of Social Robotics*, 5(4):441–456.

[Ezell et al., 2012] Ezell, S., Field, T., Nadel, J., Newton, R., Murrey, G., Siddalingappa, V., Allender, S., Grace, A., et al. (2012). Imitation effects on joint attention behaviors of children with autism. *Psychology*, 3(09):681.

[Farr et al., 2010] Farr, W., Yuill, N., and Raffle, H. (2010). Social benefits of a tangible user interface for children with autistic spectrum conditions. *Autism*.

[Feil-Seifer and Mataric, 2005] Feil-Seifer, D. and Mataric, M. J. (2005). Defining socially assistive robotics. In *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*, pages 465–468. IEEE.

[Ferguson et al., 2005] Ferguson, H., Myles, B. S., and Hagiwara, T. (2005). Using a personal digital assistant to enhance the independence of an adolescent with asperger syndrome. *Education and Training in Developmental Disabilities*, pages 60–67.

[Ferrari et al., 2009] Ferrari, E., Robins, B., and Dautenhahn, K. (2009). Therapeutic and educational objectives in robot assisted play for children with autism. In *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pages 108–114. IEEE.

[Ferrell et al., 1987] Ferrell, W., Gandevia, S., and McCloskey, D. (1987). The role of joint receptors in human kinaesthesia when intramuscular receptors cannot contribute. *The Journal of physiology*, 386(1):63–71.

[Field et al., 2001] Field, T., Field, T., Sanders, C., and Nadel, J. (2001). Children with autism display more social behaviors after repeated imitation sessions. *Autism*, 5(3):317–323.

[Fluckiger and Baumberger, 2003] Fluckiger, M. and Baumberger, B. (2003). Estimation of a texture velocity projected on the ground surface. *Investigative Ophthalmology & Visual Science*, 44(13):4087–4087.

[Foxx, 2008] Foxx, R. M. (2008). Applied behavior analysis treatment of autism: The state of the art. *Child and adolescent psychiatric clinics of North America*, 17(4):821–834.

[Froehlich et al., 2012] Froehlich, A., Anderson, J., Bigler, E., Miller, J., Lange, N., DuBray, M., Cooperrider, J., Cariello, A., Nielsen, J., and Lainhart, J. E. (2012). Intact prototype formation but impaired generalization in autism. *Research in autism spectrum disorders*, 6(2):921–930.

[García-Primo et al., 2014] García-Primo, P., Hellendoorn, A., Charman, T., Roeyers, H., Dereu, M., Roge, B., Baduel, S., Muratori, F., Narzisi, A., Van Daalen, E., et al. (2014). Screening for autism spectrum disorders: state of the art in europe. *European child & adolescent psychiatry*, 23(11):1005–1021.

[Gepner et al., 1995] Gepner, B., Mestre, D., Masson, G., and de Schonen, S. (1995). Postural effects of motion vision in young autistic children. *NeuroReport*, 6(8):1211–1214.

[Gepner and Mestre, 2002] Gepner, B. and Mestre, D. R. (2002). Brief report: postural reactivity to fast visual motion differentiates autistic from children with asperger syndrome. *Journal of Autism and Developmental disorders*, 32(3):231–238.

[Golan et al., 2010] Golan, O., Ashwin, E., Granader, Y., McClintock, S., Day, K., Leggett, V., and Baron-Cohen, S. (2010). Enhancing emotion recognition in children with autism spectrum conditions: An intervention using animated vehicles with real emotional faces. *Journal of autism and developmental disorders*, 40(3):269–279.

[Goodwin et al., 1972] Goodwin, G. M., McCloskey, D. I., and Matthews, P. B. (1972). Proprioceptive illusions induced by muscle vibration: contribution by muscle spindles to perception? *Science*, 175(4028):1382–1384.

[Gowen and Hamilton, 2013] Gowen, E. and Hamilton, A. (2013). Motor abilities in autism: a review using a computational context. *Journal of autism and developmental disorders*, 43(2):323–344.

[Greffou et al., 2012] Greffou, S., Bertone, A., Hahler, E.-M., Hanssens, J.-M., Mottron, L., and Faubert, J. (2012). Postural hypo-reactivity in autism is contingent on development and visual environment: a fully immersive virtual reality study. *Journal of autism and developmental disorders*, 42(6):961–970.

[Greffou et al., 2008] Greffou, S., Bertone, A., Hanssens, J.-M., and Faubert, J. (2008). Development of visually driven postural reactivity: a fully immersive virtual reality study. *Journal of Vision*, 8(11):15–15.

[Grossard and Grynszpan, 2015] Grossard, C. and Grynszpan, O. (2015). Entraînement des compétences assistées par les technologies numériques dans l'autisme: une revue. *Enfance*, 2015(01):67–85.

[Hart, 2005] Hart, M. (2005). Autism/excel study. In *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility*, pages 136–141. ACM.

[Haswell et al., 2009] Haswell, C. C., Izawa, J., Dowell, L. R., Mostofsky, S. H., and Shadmehr, R. (2009). Representation of internal models of action in the autistic brain. *Nature neuroscience*, 12(8):970–972.

[Herbert et al., 2004] Herbert, M. R., Ziegler, D. A., Makris, N., Filipek, P. A., Kemper, T. L., Normandin, J. J., Sanders, H. A., Kennedy, D. N., and Caviness, V. S. (2004). Localization of white matter volume increase in autism and developmental language disorder. *Annals of neurology*, 55(4):530–540.

[Hilton et al., 2010] Hilton, C. L., Harper, J. D., Kueker, R. H., Lang, A. R., Abbacchi, A. M., Todorov, A., and LaVesser, P. D. (2010). Sensory responsiveness as a predictor of social severity in children with high functioning autism spectrum disorders. *Journal of autism and developmental disorders*, 40(8):937–945.

[Isableu et al., 2011] Isableu, B., Fourre, B., Vuillerme, N., Giraudet, G., and Amorim, M.-A. (2011). Differential integration of visual and kinaesthetic signals to upright stance. *Experimental brain research*, 212(1):33–46.

[Isableu et al., 2010] Isableu, B., Ohlmann, T., Cremieux, J., Vuillerme, N., Amblard, B., and Gresty, M. (2010). Individual differences in the ability to identify, select and use appropriate frames of reference for perceptuo-motor control. *Neuroscience*, 169(3):1199–1215.

[Isableu and Vuillerme, 2006] Isableu, B. and Vuillerme, N. (2006). Differential integration of kinaesthetic signals to postural control. *Experimental brain research*, 174(4):763–768.

[Izawa et al., 2012] Izawa, J., Pekny, S. E., Marko, M. K., Haswell, C. C., Shadmehr, R., and Mostofsky, S. H. (2012). Motor learning relies on integrated sensory inputs in adhd, but over-selectively on proprioception in autism spectrum conditions. *Autism Research*, 5(2):124–136.

[Johnson et al., 2007] Johnson, C. P., Myers, S. M., et al. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics*, 120(5):1183–1215.

[Jones et al., 2009] Jones, C. R., Happé, F., Baird, G., Simonoff, E., Marsden, A. J., Tregay, J., Phillips, R. J., Goswami, U., Thomson, J. M., and Charman, T. (2009). Auditory discrimination and auditory sensory behaviours in autism spectrum disorders. *Neuropsychologia*, 47(13):2850–2858.

[Jung et al., 2006] Jung, K.-E., Lee, H.-J., Lee, Y.-S., Cheong, S.-S., Choi, M.-Y., Suh, D.-S., Suh, D., Oah, S., Lee, S., and Lee, J.-H. (2006). The application of a sensory integration treatment based on virtual reality-tangible interaction for children with autistic spectrum disorder. *PsychNology Journal*, 4(2):145–159.

[Kaiser et al., 2016] Kaiser, M. D., Yang, D. Y.-J., Voos, A. C., Bennett, R. H., Gordon, I., Pretzsch, C., Beam, D., Keifer, C., Eilbott, J., McGlone, F., et al. (2016). Brain mechanisms for processing affective (and nonaffective) touch are atypical in autism. *Cerebral Cortex*, 26(6):2705–2714.

[Kelley, 1983] Kelley, J. F. (1983). An empirical methodology for writing user-friendly natural language computer applications. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 193–196. ACM.

[Kim et al., 2013] Kim, E. S., Berkovits, L. D., Bernier, E. P., Leyzberg, D., Shic, F., Paul, R., and Scassellati, B. (2013). Social robots as embedded reinforcers of social behavior in children with autism. *Journal of autism and developmental disorders*, 43(5):1038–1049.

[Kleinsmith and Bianchi-Berthouze, 2013] Kleinsmith, A. and Bianchi-Berthouze, N. (2013). Affective body expression perception and recognition: A survey. *IEEE Transactions on Affective Computing*, 4(1):15–33.

[Kleinsmith et al., 2006] Kleinsmith, A., De Silva, P. R., and Bianchi-Berthouze, N. (2006). Cross-cultural differences in recognizing affect from body posture. *Interacting with Computers*, 18(6):1371–1389.

[Klin et al., 2009] Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., and Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*, 459(7244):257–261.

[Kluzik et al., 2007] Kluzik, J., Peterka, R. J., and Horak, F. B. (2007). Adaptation of postural orientation to changes in surface inclination. *Experimental brain research*, 178(1):1–17.

[Kohen-Raz et al., 1992] Kohen-Raz, R., Volkman, F. R., and Cohen, D. J. (1992). Postural control in children with autism. *Journal of Autism and Developmental Disorders*, 22(3):419–432.

[Kopp et al., 2006] Kopp, S., Krenn, B., Marsella, S., Marshall, A. N., Pelachaud, C., Pirker, H., Thórisson, K. R., and Vilhjálmsson, H. (2006). Towards a common framework for multimodal generation: The behavior markup language. In *International Workshop on Intelligent Virtual Agents*, pages 205–217. Springer.

[Kozima et al., 2005] Kozima, H., Nakagawa, C., and Yasuda, Y. (2005). Interactive robots for communication-care: A case-study in autism therapy. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*, pages 341–346. IEEE.

[Kugiumutzakis, 1998] Kugiumutzakis, G. (1998). Neonatal imitation in the intersubjective companion space. *Intersubjective communication and emotion in early ontogeny*, pages 63–88.

[Lei et al., 2015] Lei, J., Song, M., Li, Z.-N., and Chen, C. (2015). Whole-body humanoid robot imitation with pose similarity evaluation. *Signal Processing*, 108:136–146.

[Liu and Chepyator-Thomson, 2009] Liu, W. and Chepyator-Thomson, J. R. (2009). Field dependence–independence and physical activity engagement among middle school students. *Physical Education and Sport Pedagogy*, 14(2):125–136.

[Lord et al., 2000] Lord, C., Risi, S., Lambrecht, L., Cook Jr, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., and Rutter, M. (2000). The autism diagnostic observation schedule—generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of autism and developmental disorders*, 30(3):205–223.

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

[Luteijn et al., 2000] Luteijn, E., Luteijn, F., Jackson, S., Volkmar, F., and Minderaa, R. (2000). The children's social behavior questionnaire for milder variants of pdd problems: Evaluation of the psychometric characteristics. *Journal of autism and developmental disorders*, 30(4):317–330.

[Matarić and Scassellati, 2016] Matarić, M. J. and Scassellati, B. (2016). Socially assistive robotics. In *Springer Handbook of Robotics*, pages 1973–1994. Springer.

[Meeren et al., 2005] Meeren, H. K., van Heijnsbergen, C. C., and de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences of the United States of America*, 102(45):16518–16523.

[Mehrabian, 1972] Mehrabian, A. (1972). *Nonverbal communication*. Transaction Publishers.

[Meltzoff et al., 2010] Meltzoff, A. N., Brooks, R., Shon, A. P., and Rao, R. P. (2010). "social" robots are psychological agents for infants: A test of gaze following. *Neural Networks*, 23(8):966–972.

[Miller et al., 2014] Miller, M., Chukoskie, L., Zinni, M., Townsend, J., and Trauner, D. (2014). Dyspraxia, motor function and visual–motor integration in autism. *Behavioural brain research*, 269:95–102.

[Milne, 2011] Milne, E. (2011). Increased intra-participant variability in children with autistic spectrum disorders: evidence from single-trial analysis of evoked eeg. *Frontiers in psychology*, 2:51.

[Molloy et al., 2003] Molloy, C. A., Dietrich, K. N., and Bhattacharya, A. (2003). Postural stability in children with autism spectrum disorder. *Journal of autism and developmental disorders*, 33(6):643–652.

[Mundy and Newell, 2007] Mundy, P. and Newell, L. (2007). Attention, joint attention, and social cognition. *Current directions in psychological science*, 16(5):269–274.

[Myers et al., 2007] Myers, S. M., Johnson, C. P., et al. (2007). Management of children with autism spectrum disorders. *Pediatrics*, 120(5):1162–1182.

[Nadel, 1986] Nadel, J. (1986). *Imitation et communication entre jeunes enfants*, volume 13. Presses Universitaires de France-PUF.

[Nadel, 2002] Nadel, J. (2002). Imitation and imitation recognition: Functional use in preverbal infants and nonverbal children with autism. *The imitative mind: Development, evolution, and brain bases*, pages 42–62.

[Nadel, 2004] Nadel, J. (2004). Do children with autism understand imitation as intentional interaction? *Journal of Cognitive & Behavioral Psychotherapies*, 4(2).

[Nadel, 2014] Nadel, J. (2014). *How imitation boosts development: in infancy and autism spectrum disorder*. OUP Oxford.

[Nadel and Butterworth, 1999] Nadel, J. and Butterworth, G. (1999). *Imitation in infancy*, volume 16. Cambridge University Press Cambridge.

[Nadel-Brulfert and Baudonniere, 1982] Nadel-Brulfert, J. and Baudonniere, P. (1982). The social function of reciprocal imitation in 2-year-old peers. *International Journal of Behavioral Development*, 5(1):95–109.

[Noris et al., 2012] Noris, B., Nadel, J., Barker, M., Hadjikhani, N., and Billard, A. (2012). Investigating gaze of children with asd in naturalistic settings. *PloS one*, 7(9):e44144.

[Organization et al., 2013] Organization, W. H. et al. (2013). Meeting report: autism spectrum disorders and other developmental disorders: from raising awareness to building capacity: World health organization, geneva, switzerland 16-18 september 2013.

[Panerai et al., 2002] Panerai, S., Ferrante, L., and Zingale, M. (2002). Benefits of the treatment and education of autistic and communication handicapped children (teacch) programme as compared with a non-specific approach. *Journal of intellectual disability research*, 46(4):318–327.

[Park and Schowengerdt, 1982] Park, S. K. and Schowengerdt, R. A. (1982). Image sampling, reconstruction, and the effect of sample-scene phasing. *Applied Optics*, 21(17):3142–3151.

[Patwardhan and Knapp, 2016] Patwardhan, A. and Knapp, G. (2016). Multimodal affect recognition using kinect. *arXiv preprint arXiv:1607.02652*.

[Pop et al., 2013] Pop, C. A., Simut, R., Pintea, S., Saldien, J., Rusu, A., David, D., Vanderfaeillie, J., Lefeber, D., and Vanderborght, B. (2013). Can the social robot probo help children with autism to identify situation-based emotions? a series of single case experiments. *International Journal of Humanoid Robotics*, 10(03):1350025.

[Ranatunga et al., 2013] Ranatunga, I., Beltran, M., Torres, N. A., Bugnariu, N., Patterson, R. M., Garver, C., and Popa, D. O. (2013). Human-robot upper body gesture imitation analysis for autism spectrum disorders. In *International Conference on Social Robotics*, pages 218–228. Springer.

[Ravaioli et al., 2005] Ravaioli, E., Oie, K. S., Kiemel, T., Chiari, L., and Jeka, J. J. (2005). Nonlinear postural control in response to visual translation. *Experimental brain research*, 160(4):450–459.

[Robertson and Simmons, 2013] Robertson, A. E. and Simmons, D. R. (2013). The relationship between sensory sensitivity and autistic traits in the general population. *Journal of Autism and Developmental disorders*, 43(4):775–784.

[Robins et al., 2004] Robins, B., Dickerson, P., Stribling, P., and Dautenhahn, K. (2004). Robot-mediated joint attention in children with autism: A case study in robot-human interaction. *Interaction studies*, 5(2):161–198.

[Rogers and Ozonoff, 2005] Rogers, S. J. and Ozonoff, S. (2005). Annotation: What do we know about sensory dysfunction in autism? a critical review of the empirical evidence. *Journal of Child Psychology and Psychiatry*, 46(12):1255–1268.

[Rogers and Pennington, 1991] Rogers, S. J. and Pennington, B. F. (1991). A theoretical approach to the deficits in infantile autism. *Development and psychopathology*, 3(2):137–162.

[Roll and Vedel, 1982] Roll, J. and Vedel, J. (1982). Kinaesthetic role of muscle afferents in man, studied by tendon vibration and microneurography. *Experimental Brain Research*, 47(2):177–190.

[Rump et al., 2009] Rump, K. M., Giovannelli, J. L., Minshew, N. J., and Strauss, M. S. (2009). The development of emotion recognition in individuals with autism. *Child development*, 80(5):1434–1447.

[Russell, 1994] Russell, J. A. (1994). Is there universal recognition of emotion from facial expressions? a review of the cross-cultural studies. *Psychological bulletin*, 115(1):102.

[Rutter et al., 2003] Rutter, M., Bailey, A., and Lord, C. (2003). *The social communication questionnaire: Manual*. Western Psychological Services.

[Rutter et al., 2012] Rutter, M., DiLavore, P., Risi, S., Gotham, K., and Bishop, S. (2012). Autism diagnostic observation schedule: Ados-2. *Torrance, CA: Western Psychological Services*.

[Salowitz et al., 2013] Salowitz, N. M., Eccarius, P., Karst, J., Carson, A., Schohl, K., Stevens, S., Van Hecke, A. V., and Scheidt, R. A. (2013). Brief report: visuo-spatial guidance of movement during gesture imitation and mirror drawing in children with autism spectrum disorders. *Journal of autism and developmental disorders*, 43(4):985–995.

[Salter et al., 2010] Salter, T., Michaud, F., and Larouche, H. (2010). How wild is wild? a taxonomy to characterize the ‘wildness’ of child-robot interaction. *International Journal of Social Robotics*, 2(4):405–415.

[Saracho, 2003] Saracho, O. (2003). Matching teachers' and students' cognitive styles. *Early Child Development and Care*, 173(2-3):161–173.

[Saragih et al., 2009] Saragih, J. M., Lucey, S., and Cohn, J. F. (2009). Face alignment through subspace constrained mean-shifts. In *2009 IEEE 12th International Conference on Computer Vision*, pages 1034–1041. IEEE.

[Scassellati et al., 2012] Scassellati, B., Admoni, H., and Mataric, M. (2012). Robots for use in autism research. *Annual Review of Biomedical Engineering*, 14:275–294.

[Scherer, 1995] Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of voice*, 9(3):235–248.

[Scherer et al., 2012] Scherer, S., Marsella, S., Stratou, G., Xu, Y., Morbini, F., Egan, A., Morency, L.-P., et al. (2012). Perception markup language: Towards a standardized representation of perceived nonverbal behaviors. In *International Conference on Intelligent Virtual Agents*, pages 455–463. Springer.

[Schmuckler, 1997] Schmuckler, M. A. (1997). Children's postural sway in response to low- and high-frequency visual information for oscillation. *Journal of Experimental Psychology: Human Perception and Performance*, 23(2):528.

[Seibert and Hogan, 1982] Seibert, J. M. and Hogan, A. E. (1982). *Procedures Manual for the Early Social-Communication Scales (ESCS).* Mailman Center for Child Development, University of Miami.

[Seltzer et al., 2004] Seltzer, M. M., Shattuck, P., Abbeduto, L., and Greenberg, J. S. (2004). Trajectory of development in adolescents and adults with autism. *Mental retardation and developmental disabilities research reviews*, 10(4):234–247.

[Shattuck et al., 2007] Shattuck, P. T., Seltzer, M. M., Greenberg, J. S., Orsmond, G. I., Bolt, D., Kring, S., Lounds, J., and Lord, C. (2007). Change in autism symptoms and maladaptive behaviors in adolescents and adults with an autism spectrum disorder. *Journal of autism and developmental disorders*, 37(9):1735–1747.

[Silver and Oakes, 2001] Silver, M. and Oakes, P. (2001). Evaluation of a new computer intervention to teach people with autism or asperger syndrome to recognize and predict emotions in others. *Autism*, 5(3):299–316.

[Simmons et al., 2009] Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., and Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision research*, 49(22):2705–2739.

[Skuse et al., 2004] Skuse, D., Warrington, R., Bishop, D., Chowdhury, U., Lau, J., Mandy, W., and Place, M. (2004). The developmental, dimensional and diagnostic interview (3di): A novel computerized assessment for autism spectrum disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, 43(5):548–558.

[Srinivasan et al., 2015] Srinivasan, S. M., Kaur, M., Park, I. K., Gifford, T. D., Marsh, K. L., and Bhat, A. N. (2015). The effects of rhythm and robotic interventions on the imitation/praxis, interpersonal synchrony, and motor performance of children with autism spectrum disorder (asd): A pilot randomized controlled trial. *Autism research and treatment*, 2015.

[Taheri et al., 2015] Taheri, A., Alemi, M., Meghdari, A., Pouretemad, H., and Holderread, S. (2015). Clinical application of humanoid robots in playing imitation games for autistic children in iran. *Procedia-Social and Behavioral Sciences*, 176:898–906.

[Tapus et al., 2007] Tapus, A., Mataric, M., and Scassellati, B. (2007). Attention to facial emotion expressions in children with autism. *IEEE Robotics and Automation Magazine (RAM), Special Issue on Grand Challenges in Robotics*, 14(1).

[Tapus et al., 2012] Tapus, A., Peca, A., Aly, A., Pop, C., Jisa, L., Pintea, S., Rusu, A. S., and David, D. O. (2012). Children with autism social engagement in interaction with nao, an imitative robot–a series of single case experiments. *Interaction studies*, 13(3):315–347.

[Tapus et al., 2008] Tapus, A., Ţăpuş, C., and Matarić, M. J. (2008). User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics*, 1(2):169–183.

[Thill et al., 2012] Thill, S., Pop, C. A., Belpaeme, T., Ziemke, T., and Vanderborght, B. (2012). Robot-assisted therapy for autism spectrum disorders with (partially) autonomous control: Challenges and outlook. *Paladyn, Journal of Behavioral Robotics*, 3(4):209–217.

[Tomasello, 2009] Tomasello, M. (2009). *The cultural origins of human cognition*. Harvard university press.

[Torres et al., 2013] Torres, E. B., Brincker, M., Isenhower, R. W., Yanovich, P., Stigler, K. A., Nurnberger, J. I., Metaxas, D. N., and José, J. V. (2013). Autism: the micro-movement perspective.

[Uljarevic and Hamilton, 2013] Uljarevic, M. and Hamilton, A. (2013). Recognition of emotions in autism: a formal meta-analysis. *Journal of autism and developmental disorders*, 43(7):1517–1526.

[Van Asten et al., 1988] Van Asten, W., Gielen, C., and van der Gon, J. D. (1988). Postural adjustments induced by simulated motion of differently structured environments. *Experimental Brain Research*, 73(2):371–383.

[Vanderborght et al., 2012] Vanderborght, B., Simut, R., Saldien, J., Pop, C., Rusu, A. S., Pintea, S., Lefeber, D., and David, D. O. (2012). Using the social robot probo as a social story telling agent for children with asd. *Interaction Studies*, 13(3):348–372.

[Vanvuchelen et al., 2007] Vanvuchelen, M., Roeyers, H., and De Weerdt, W. (2007). Nature of motor imitation problems in school-aged boys with autism a motor or a cognitive problem? *Autism*, 11(3):225–240.

[Vaufreydaz et al., 2016] Vaufreydaz, D., Johal, W., and Combe, C. (2016). Starting engagement detection towards a companion robot using multimodal features. *Robotics and Autonomous Systems*, 75:4–16.

[Voos et al., 2013] Voos, A. C., Pelphrey, K. A., and Kaiser, M. D. (2013). Autistic traits are associated with diminished neural response to affective touch. *Social cognitive and affective neuroscience*, 8(4):378–386.

[Wainer et al., 2010] Wainer, J., Dautenhahn, K., Robins, B., and Amirabdollahian, F. (2010). Collaborating with kaspar: Using an autonomous humanoid robot to foster cooperative dyadic play among children with autism. pages 631–638.

[Wallbott, 1998] Wallbott, H. G. (1998). Bodily expression of emotion. *European journal of social psychology*, 28(6):879–896.

[Walter, 2007] Walter, E. L. (2007). *Visuospatial contextual processing: Illusions, hidden figures and autistics traits*. University of Oregon.

[Warren et al., 2015] Warren, Z. E., Zheng, Z., Swanson, A. R., Bekele, E., Zhang, L., Crittendon, J. A., Weitlauf, A. F., and Sarkar, N. (2015). Can robotic interaction improve joint attention skills? *Journal of autism and developmental disorders*, 45(11):3726–3734.

[Williams et al., 2004] Williams, J. H., Whiten, A., and Singh, T. (2004). A systematic review of action imitation in autistic spectrum disorder. *Journal of autism and developmental disorders*, 34(3):285–299.

[Yoder et al., 2006] Yoder, P. J., McDuffie, A. S., Charman, T., and Stone, W. (2006). Treatment of responding to and initiating joint attention. *Social and communication development in autism spectrum disorders: Early identification, diagnosis, and intervention*, pages 117–142.

Pauline Chevalier

ENSTA-ParisTech
828, Bld des Maréchaux, 91120 Palaiseau FRANCE
+33 6 74 73 95 53
pauline.chevalier@ensta-paristech.fr
http://perso.ensta-paristech.fr/~chevalier

Education

2013: **M.S., Electronic Engineering**
Université Pierre et Marie Curie, Paris, France
Master Sciences de l'Ingnieur Electronique: Orientation Systme Avancé et Robotique - Spécialité Systèmes Intelligents et Robotique

2012: **B.S. with Honors, Electronic Engineering**
Université Pierre et Marie Curie, Paris, France
Master Sciences de l'Ingénieur Electronique: Orientation Informatique Industrielle, Image, Signal et Robotique

2011: **B.S. Science and Technology, Electronic Engineering**
Université Pierre et Marie Curie, Paris, France

Research Experience

Dec 2013 - present: PhD Student, Robotics and Computer Vision Lab, ENSTA-ParisTech, 828, Bld des Marchaux, 91120 Palaiseau, Supervisors: Prof. Adriana Tapus, Prof. Jean-Claude Martin, and Prof. Brice Isableu.
Proprioceptive, Kinematic, Socially and Individualized Human-Machine Interaction.

Feb 2013 - Aug 2013: Internship, Institut des Systèmes Intelligents et de Robotique (ISIR), 5 Place Jussieu, 75005 Paris, Supervisor : Kevin Bailly, *(6 months)*
Facial expression analysis for Human-Virtual Agent Interaction.

May 2012 - Jul 2012: Internship, Institut des Systèmes Intelligents et de Robotique (ISIR), 5 Place Jussieu, 75005 Paris, Supervisor: Kevin Bailly, *(2 months)*
Detection and visualization of facial expression by vision.

Teaching Experience

Jan 2015 - Apr 2015: Teaching Assistant for Projet tutoré - Compléments (Project with Nao robot), Second year of I.U.T. d'Informatique, 27 hours, I.U.T. d'Informatique d'Orsay, Orsay, France.

Sep 2015 - Oct 2015: Teaching Assistant for Méthodologie de la production de logiciel, Second year of I.U.T. d'Informatique, 10.5 hours, I.U.T. d'Informatique d'Orsay, Orsay, France.

Sep 2015 - Jan 2015: Teaching Assistant for Conception de documents et d'interfaces numeriques, First year of I.U.T. d'Informatique, 25.5 hours, I.U.T. d'Informatique d'Orsay, Orsay, France.

Mar 2015 - Jun 2015: Teaching Assistant for Computer Science Project - Picross (*Projet Informatique - Picross*), First year of Engineering School, 24 hours, ENSTA-ParisTech, Palaiseau, France.

Sep 2014 - Dec 2014: Teaching Assistant for Algorithmic and Programming - in Python (*Algorithmique et Programmation - en Python*), First year of Engineering School, 21.25 hours, ENSTA-ParisTech, Palaiseau, France.

Honors and Awards

2016 - Travel Fellowhip "International Exchanges for PhD Students", Interface Doctoral School 2016
2016 - Nominated for Prix Innovation jeune chercheur.se, Handiversité 2016
2015 - Prix Doctorant Demenÿ-Vaucansson 2015
2015 - Best Student Paper Award, RAAD 2015
2013 - Ile de France Doctoral Fellowship, HANDICAP theme

Publications

Refereed Journal Articles

1. **Chevalier, P.**, Iacob, D.-O., Martin, J.-C., Isableu, B., Bazile, C., Tapus, A., Joint Attention in Human-Robot Interaction: Impact of Sensory Preferences of Individuals with Autism, International Journal of Social Robotics (under review summer 2016)
2. **Chevalier, P.**, Martin, J.-C., Isableu, B., Bazile, C., Tapus, A., Impact of Sensory Preferences of Autistic Users on the Recognition of Emotions Expressed by two Robots, an Avatar and a Human, Autonomous Robots, Special Issue on Assistive and Rehabilitation Robotics, May 2016

Refereed Book Chapter Papers

3. **Chevalier, P.**, Martin, J.-C., Isableu, B., Bazile, C., Tapus, A., Proprioceptive and Kinematic Profiles for Customized Human-Robot Interaction for People Suffering from Autism, Autism - Paradigms and Clinical Application (under review summer 2016)
4. **Chevalier, P.**, Martin, J.-C., Isableu, B., Tapus, A., Individuals with Autism: Analysis of the First Interaction with Nao Robot Based on Their Proprioceptive and Kinematic Profiles, In Advances in Robot Design and Intelligent Control, 225-233.

Refereed Conference Papers

5. **Chevalier, P.**, Raiola, G., Martin, J.-C., Isableu, B., Bazile, C., Tapus, A., Do Sensory Preferences of Children with Autism Impact an Imitation Task with a Robot?, Under Review to HRI2017
6. Sanchez Restrepo, S., Raiola, G., **Chevalier, P.**, Lamy, X., Sidobre, D., Iterative Virtual Guides Programming for Comanipulation Robots, Under Review for The International Conference on Robotics and Automation (ICRA2017).
7. **Chevalier, P.**, Martin, J.-C., Isableu, B., Bazile, C., Tapus, A., Joint Attention with a Robot: Impact of Sensory Preferences of Children with Autism, In Proceedings of RO-MAN 2016, New York, August 2016
8. **Chevalier, P.**, Martin, J.-C., Isableu, B., Tapus, A., Impact of Personality on the Recognition of Emotion Expressed via Human, Virtual and Robotic Embodiments, In Proceedings of RO-MAN 2015, Kobe, Japan, September 2015

Refereed Poster Papers

9. **Chevalier, P.**, Martin, J.-C., Isableu, B., Bazile, C., Tapus, A., Human-Robot Interaction: Impact of Sensory Preferences for Users with Autistic Spectrum Disorder, MB&E 2016 / 21 - 23 November, Lille
10. **Chevalier, P.**, Martin, J.-C., Isableu, B., Bazile, C., Tapus, A., Attention Conjointe avec un Robot : Impact des Prfrences Sensorielles chez des Individus avec Autisme, Handiversité 2016
11. **Chevalier, P.**, Martin, J.-C., Isableu, B., Bazile, C., Tapus, A., Attention Conjointe avec un Robot : Impact des Prfrences Sensorielles chez des Individus avec Autisme, Handiversité 2016
12. **Chevalier, P.**, Martin, J.-C., Isableu, B., Bazile, C., Tapus, A., Social Personalized Human-Machine Interaction for People with Autism: A close look at Proprioceptive and Visual Orientation Integration, International Meeting for Autism Research (IMFAR2015), Salt Lake City, Utah, USA, May 2015
13. **Chevalier, P.**, Martin, J.-C., Isableu, B., Bazile, C., Tapus, A., Social Personalized Human-Machine Interaction for People with Autism: Defining User Profiles and First Contact with a Robot, In Proceedings of IEEE Human-Robot Interaction Conference (HRI2015) Late Breaking Report, Portland, Oregon, USA, March 2015

Refereed Workshop and Symposia Papers

14. **Chevalier, P.**, Raiola, G., Isableu, B., Martin, J.-C., Bazile, C., and Tapus, A., Impact des profils sensoriels sur l'imitation des mouvements d'un robot par des participants autistes, Journée FéDeV 2016
15. **Chevalier, P.**, Isableu, B., Martin, J.-C., and Tapus, A., Profilage Visuel et Proprioceptif pour l'Elaboration d'Interactions Homme-Robot Adaptées, Journée FéDeV 2015
16. **Chevalier, P.**, Isableu, B., Martin, J.-C., and Tapus, A., Evaluation of Child-Robot Interaction: why is it different from adults?, 1st Workshop on Evaluating Child-Robot Interaction at ICSR, Paris, October 26 2015
17. **Chevalier, P.**, Social Personalized Human-Machine Interaction for People with Autism, In Proceedings of IEEE Human-Robot Interaction Conference (HRI) Pioneers Workshop, Portland, Oregon, USA, March 2015
18. **Chevalier, P.**, Martin, J.-C., Isableu, B., Bazile, C., Tapus, A., Social Personalized Human-Robot Interaction for People with Autism: A close look at Proprioceptive and Visual Orientation Integration, Journées Nationales de la Robotique Interactive 2014 (JNRI-2014), Toulouse, November, 2014

Other

19. Lanfranchi, S., and **Chevalier, P.**, "La Recherche et la Robotique, Deux Alliées de Choix dans le Combat contre l'Autisme", In ENSTActualités 14, Printemps 2015

Service

Conferences Organizing Committees

Local Organising Committee, International Conference on Social Robotics (ICSR) 2015

Reviewer for Profesional Journals and Conferences

International Journal of Social Robotics
Interaction Studies
IEEE Robotics and Automation Letters (RA-L)
ACM/IEEE International Conference on Human-Robot Interaction (HRI)
IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)
International Conference of Social Robotics (ICSR)
IEEE Transactions on Cognitive and Developmental Systems
IEEE International Conference on Robotics and Automation (ICRA)

Invited Talks

Invited Talks - excluding conference paper talk

"Human-Robot Interaction: Impact of Sensory Preferences of Individuals with Autism", Guest Lecture, at the Social Robotic Lab, Yale University, September 2016

Others

Participation to Science Festival (*Fête de la Science*) at ENSTA-ParisTech in 2014 and 2015, and at Université Paris-Sud in 2015 and 2016.

Handicap Day (*Journée du Handicap*), ENSTA-ParisTech, September the 29th, 2016

Co-organizer of PhD students Meetings in ENSTA-ParisTech U2IS laboratory, winter and spring 2016

ÉCOLE DOCTORALE
INTERFACES
Approches Interdisciplinaires : Fondements, Applications et Innovation

Titre : Impact des préférences sensorielles chez les individus souffrant de troubles du spectre autistique sur leur interaction sociale avec un robot

Mots clés : Interaction Homme-Machine; Interaction Sociale; Communication non-verbale; Autisme; Proprioception; Vision

Résumé : L'objectif de ce travail de thèse est de permettre sur le long terme de proposer des interactions sociales personnalisées exploitant l'attirance que les personnes souffrant de Troubles du Spectre Autistique (TSA) ont envers les robots humanoïdes, tels que Nao (Softbank Robotics), afin d'améliorer leurs capacités d'interaction sociale. Trois institutions spécialisées pour personnes atteintes de TSA participent à notre étude : l'IME Notre Ecole et l'IME MAIA, Institut médico-éducatif pour enfants et adolescents atteints de TSA et la Lendemaine, Foyer d'Aide Médicalisée pour adultes atteints de TSA.

Les différences inter-individuelles sont très présentes dans les TSA et elles impactent différemment le comportement de chaque individu avec TSA dans sa vie (par exemples la communication, l'attention jointe, ou encore les troubles moteurs, à des degrés différents pour chaque individu), et dans cette étude, durant leur interaction sociale avec un robot.

Afin d'envisager à long terme une interaction personnalisée pour chaque participant, une première étape a consisté à définir leur profil sensorimoteur. L'hypothèse qui guide notre étude est que l'intégration des informations visuelles et proprioceptives (perception, consciente ou non, de la position et des changements des différentes parties du corps) d'une personne joue un rôle sur ses capacités sociales. Une personne qui réagit peu aux informations visuelles et qui utilise les informations proprioceptives relatives à ses mouvements ou à la position de son corps de manière exacerbée, aurait plus de difficultés à s'engager et à maintenir une interaction sociale.

Les profils sensoriels des participants ont été évalués à l'aide du test du Profil Sensoriel de Dunn et du test sensorimoteur impliquant une scène mobile virtuelle afin d'évaluer leur dépendance visuelle et proprioceptive. L'analyse des données a permis de classer nos participants en trois groupes montrant des comportements différents face aux informations proprioceptives et visuelles, et à leur intégration.

Nous avons ensuite étudié les liens entre les profils sensoriels des participants et leurs différents comportements sociaux à travers plusieurs tâches impliquées dans les interactions sociales : (1) reconnaissance d'émotions exprimées par deux robots, un avatar et une personne ; (2) interaction sociale avec le robot Nao sur la salutation ; (3) attention conjointe avec le robot Nao, et (4) imitation avec le robot Nao. Cette dernière tâche a fait l'objet de sessions répétées sur huit semaines (modèle de thérapie sur l'apprentissage et de renforcement de l'imitation pour enfants avec TSA).

A travers ces études, nous avons pu observer que les participants ayant une plus forte dépendance à la proprioception et une indépendance au champ visuel ont eu plus de difficultés à interagir avec le robot (moins de regards vers le robot, moins de réponses à l'attention conjointe, plus de difficultés à reconnaitre les émotions, et à imiter un partenaire) que les autres participants.

Nous avons pu observer que les sessions avec le robot Nao ont eu un effet bénéfique chez les participants avec TSA. A la suite des sessions répétées avec le robot Nao, les participants ont montré une amélioration de leurs capacités sociales (regard vers le partenaire, imitations) vers un partenaire d'imitation humain.

Ces résultats confortent l'idée d'utiliser les profils sensoriels des personnes avec TSA pour leur proposer, dans des recherches futures, des interactions personnalisées avec les robots.

Title : Impact of sensory preferences in individuals with autism spectrum disorder on their social interaction with a robot

Keywords : Human-Robot Interaction; Social Interaction; Non-verbal Communication; Autism; Proprioception; Vision

Abstract: The goal of this thesis is to provide contributions that will help in the long term to enable personalized robot-based social interaction for individuals with Autism Spectrum Disorders (ASD). This work was done in collaboration with three care facilities for people suffering from ASD: IME MAIA (France) and IME Notre Ecole, medical and educative schools for children and teenagers with ASD, and FAM La Lendemaine (France), a medical house for adults with ASD.

Inter-individual differences are present in ASD, and impact the behaviors of each individual in their lives, and in this study, during their interactions with a robot.

The first step of our work was to propose an appropriate method to define the proprioceptive and visual profiles of each of our participants. We based our work on the hypothesis that the proprioceptive (the ability of an individual to determine body segment positions (i.e., joint position sense and to detect limb movements in space) and visual integration of cues of an individual with ASD is an indicator of their social and communication skills. We posit that a mitigated behavioral response (i.e., hyporeactivity) to visual motion and an overreliance on proprioceptive information are linked in individuals with ASD to their difficulties in integrating social cues and engaging in successful social interactions. We used two methods to define the proprioceptive and visual profile of our participant: a well-known questionnaire on sensory preferences and an experimental setup. With the setup, we were able to observe three different groups of postural behaviors in our participants. Thanks to these individual profiles, we could make assumptions on the behaviors that one can expect from each of our participants during interactions with the robot.

We aimed to assess various social skills of our participants in regards to their profiles. We designed three single case studies: (1) emotion recognition with different embodiments (two robots, a virtual agent and a human); (2) a short greeting social task with the robot Nao; and (3) a game evaluating joint attention response to the robot Nao. We also conducted eight weeks-long sessions with an imitation task with Nao.

Through these studies, we were able to observe that the participants that display an overreliance on proprioceptive cues and a hyporeactivity to visual cues had more difficulties to interact with the robot (less gaze towards the robot, less answers to joint attention initiation behaviors, more difficulties to recognize emotions and to imitate a partner) than the other participants.

We were able to observe that the repeated sessions with the robot Nao were benefic for participants with ASD: after the sessions with the robot Nao, the participants showed an improvement in their social skills (gaze to the partner, imitations).

Defining such individual profiles could provide promising strategies for designing successful and adapted Human-Robot Interaction for individuals with ASD.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France