Computational Modeling of User Activity in Full-Body Interaction Environments for ASC Children:

Multimodal Analysis of Social Interaction Behaviors through Psychophysiology, System Activity, Video Coding, Questionnaires, and Body Cues

Batuhan Sayis

TESI DOCTORAL UPF / Year 2021

THESIS SUPERVISORS
Prof. Narcís Parés Burguès,
Prof. Rafael Ramírez-Meléndez,
Department of Information and Communication Technologies

upf. Universitat Pompeu Fabra Barcelona

Infinite variables of human behavior are unique colors on a palette. Without an understanding of the core theoretical and practical principles of human thought, there is no possible way for a painter to shape these colors to create his perfect masterpiece.

Acknowledgements

I am deeply grateful to my supervisors, Prof. Narcis Pares and Prof. Rafael Ramirez for their invaluable advice, continuous support, and patience during my research journey including my MSc and PhD studies. CSIM master which was co-created and coordinated by Narcis was a cornerstone in my life to apply my Computer Science knowledge in a multidisciplinary research field where I was introduced to cognitive science, psychophysiology, and interaction design. Since then, Narcis has always opened a door of opportunity to further my research interests. With my PhD study, being my main supervisor, he gave me the chance to apply my knowledge to people with special needs, to have a positive impact on their lives which was previously only limited to my personal efforts in non-governmental organizations. Narcis's clear vision, his focus, and technical inputs during designing and evaluating interaction design approaches helped me to pursue my goals. All the discussions we had, pushed my thinking and my research forward in more ways than I could possibly express here. The discussions with him were so rich that it took me a great deal of effort not to miss a single idea while taking notes. Without his constantly encouraging comments, his professional approach and, his support in connecting me with the resources I need this work would not have been possible. Thank you, Narcis, for everything.

I would also like to thank to Rafael for co-supervising my research work. Since the beginning of my research journey, he helped me to become acquainted with psychophysiology and using computational approaches to fuel my research goals. He actively supported me building interdisciplinary collaborations. He was available to answer my questions and offer guidance whenever I seek out his mentorship. His cheerful mood and taking things on the stride, are the hallmarks of Rafael that made my learning experience joyful. Thank you, Rafael, for everything.

I would also like to extend my thanks to the members of the Full-Body Interaction Lab for creating such a supportive atmosphere over the last five years. Joan, Laura, and Marie, starting my research work by following the last steps in your research was very inspiring. Ciera, thank you for your help in our collaborative research work. Olga, thank you for all the assistance you provided during the experiments and your master thesis. I am sure the PhD study now you are starting will take the lessons learned from my research to another level. Also, thank you to my other master students Claudia, Jara, Andrea, Micheal, Tan for inspiring me to become a better supervisor. Thank you to the laboratory's contributing psychologists Juan Pedro Benitez for communicating and managing tasks directly related to the special needs participants.

I gratefully acknowledge the funding received towards my PhD from Spanish Ministry of Economy and Competitiveness under the Maria de Maeztu Units of Excellence Program. I would like to thank to everyone who contributed running this program, especially to Aurelio Ruiz Garcia for always being supportive towards my research. Thank you to the clinical psychologists at the Hospital Sant Joan de Deu who lent their expertise in all aspects related to Autism and providing access to a user base which gives a pool of potential ASC participants in our research. Thank you to FAPAC which is the association that groups parent-teacher associations (AMPAs) in Barcelona for helping us accessing non-ASC participants. Thank you to Network of New Technologies in Mental Health (TECSAM) who accepted to strengthen our connections to the ASC community and centers locally and internationally. Thank you to all caregivers, families and, especially, children who have contributed to this research.

I would also like to express my sincere gratitude to Prof. Hatice Gunes, for hosting me at the Affective Intelligence & Robotics Lab at the University of Cambridge (UK). It was an enlightening experience to strengthen my research training and to discover other potential research directions. Her guidance helped me to challenge myself to think more deeply, and to explore different perspectives. In addition, I am grateful her for providing me the opportunity to be one of the volunteer coordinators of ACII conference. It was such a rewarding experience allowing me to become part of Affective Computing community from firsthand and to create potential collaboration opportunities with great scientists.

My friends, thank you to all of you for helping me get through this journey. Especially, Bora, Firat and Alp. You guys are my best friends, best examples of growing together. Bora thank you for being the real-life dictionary definition of "*rezilience*" for me. Firat thank you for always reminding me to be grateful with all the challenges that life brings ("enjoy the ride"). Alp thank you for being there literally and listening me countless of hours. And Aizea, Corinna and Patricia, thank you for all your love and support.

And most importantly my family. You have supported me in every decision I have made. My mother, thank you for being a role model and motivating me to do my best. My father, thank you for teaching me to be a curios person and feeding this curiosity whenever possible. My sister, thank you for your unconditional love and being just a second away whenever I need your support.

Abstract

Full-body Interaction experiences based on Mixed Reality (MR) systems are already playing an important role in encouraging socialization behaviors in children with Autism Spectrum Condition (ASC), as seen in the state of the art of this thesis. However, the data from these systems is multimodal in nature and complex to analyze. Fusion and analysis of this data is crucial to achieve a complete understanding of how these resources interact with each other. In this PhD Thesis, given the characteristics of full-body interaction, we developed new multimodal data gathering and evaluation techniques to better understand the effectiveness of the experience developed in our Full-body Interaction Lab (FuBInt-Lab) called Lands of Fog. This is a large-scale MR, full-body interaction environment, which allows two children to play face-to-face and explore the physical and virtual worlds simultaneously. Specifically, we developed an experimental setup for comparing Lands of Fog with a control condition based on LEGO construction toys, which includes: recording psychophysiological measures synchronized with other data sources such as observed overt behaviors and system logs of game events. In order to capture accurate psychophysiological data, we developed a wearable that is child-friendly and robust to movement artifacts in the context of ambulatory full-body interaction. In order to integrate observed overt behaviors with other data sources, we designed and developed a novel video coding protocol and an adapted coding grid conceived for Social Interaction Behaviors (SIBs) in ASC children. Using a repeated measure design, we collected data from seventy-two children (36 ASC/non-ASC dyads) from the city of Barcelona, with ages between 8-12 years old (N = 12 female, N = 60 male). Data from these trials has been organized into a public database and processed based on a semi-automatic software pipeline developed within this project. Based on this data we developed three different computational models for modelling SIBs in children with ASC during Lands of Fog sessions, compared to LEGO sessions. The results of this research support the idea that full-body interaction MR environments are capable of fostering SIBs in children with ASC with similar success as the LEGO setting, with an added advantage of being more flexible. Findings reported here shed new light on developing a tool that is mediating, guiding, and supporting the progress of the children in terms of practicing SIBs and providing structure and assistance to therapists.

Resumen

Las experiencias de interacción de cuerpo entero basadas en sistemas de Realidad Mixta (RM) juegan ya un papel importante en el fomento de conductas de socialización en niños y niñas con Trastorno del Espectro Autista (TEA), como se recoge en el estado de la cuestión de esta tesis. Sin embargo, los datos generados por estos sistemas son de naturaleza multimodal y complejos de analizar. La fusión y el análisis de estos datos son cruciales para lograr una comprensión completa de cómo estos recursos interactúan entre sí. En esta Tesis Doctoral, dadas las características de la interacción de cuerpo entero, hemos desarrollado nuevas técnicas de recolección y evaluación de datos multimodales para comprender mejor la efectividad de la experiencia desarrollada en nuestro Laboratorio de Interacción de Cuerpo Entero (FuBIntLab) llamada Lands of Fog. Esto es, un entorno de RM de gran formato, con interacción de cuerpo entero, que permite que dos niños jueguen cara a cara y puedan explorar el mundo físico y virtual simultáneamente. Específicamente, hemos desarrollado un diseño experimental para comparar Lands of Fog con una condición de control basada en los juegos de construcción LEGO, que incluye el registro de medidas psicofisiológicas sincronizadas con otras fuentes de datos, tales como comportamientos manifiestos observados y registros de los eventos del sistema del juego. Con el fin de capturar datos psicofisiológicos precisos, desarrollamos un dispositivo “wearable” apto para niños y robusto ante artefactos de movimiento en el contexto de una interacción ambulatoria de cuerpo entero. Con el fin de integrar comportamientos manifiestos observados con otras fuentes de datos, diseñamos y desarrollamos un nuevo protocolo de codificación de video y una rejilla de codificación adaptada, concebida para Comportamientos de Interacción Social (CIS) en niños TEA. Utilizando un diseño de mediciones múltiples, se recogieron datos de setenta y dos niños (36 díadas TEA / no-TEA) de la ciudad de Barcelona, con edades entre 8-12 años (N = 12 mujeres, N = 60 hombres). Los datos de estos ensayos se han organizado en una base de datos pública y se fueron procesados sobre un “pipeline” semiautomático de software desarrollado dentro de este proyecto. Basándonos en estos datos, desarrollamos tres modelos computacionales diferentes para modelar CIS en niños con TEA durante las sesiones de Lands of Fog, comparándolo con las sesiones de LEGO. Los resultados de esta investigación dan validez a la idea de que los entornos de RM de interacción de cuerpo entero son capaces de fomentar los CIS en niños con TEA con un éxito equivalente al de LEGO, con la ventaja adicional de ser más flexibles. Los hallazgos descritos aquí aportan nueva luz sobre el desarrollo de una herramienta que media, guía y apoya el progreso de los niños en la práctica de CIS, brindado estructura y ayuda a los terapeutas.

Resum

Les experiències d'interacció de cos sencer basades en sistemes de Realitat Mixta (RM) ja juguen un paper important en el foment de comportaments de socialització en nens i nenes amb Trastorn de l'Espectre Autista (TEA), com es pot observar a l'estat de la qüestió d'aquesta tesi. No obstant, les dades d'aquest sistema són de naturalesa multimodal i complexes d'analitzar. La fusió i l'anàlisi d'aquestes dades son crucials per aconseguir una comprensió completa de com aquests recursos interactuen entre sí. En aquesta Tesi Doctoral, ateses les característiques de la interacció de cos sencer, desenvolupem noves tècniques de recollida i avaluació de dades multimodals per comprendre millor l'efectivitat de l'experiència desenvolupada en el nostre Laboratori d'Interacció de Cos Sencer (FuBIntLab) anomenada Lands of Fog. En concret, un entorn de RM de gran format, amb interacció de cos sencer que permet que dos infants juguin cara a cara i puguin explorar el món físic i virtual simultàniament. Específicament, hem desenvolupat un disseny experimental per tal de comparar Lands of Fog amb una condició de control basada en jocs de construcció LEGO, que inclou: enregistrar mesures psicofisiològiques sincronitzades amb altres fonts de dades, com són el comportaments manifestos observats i els registres d'esdeveniments del sistema del joc. Amb la finalitat de capturar dades psicofisiològiques precises, hem desenvolupat un dispositiu "wearable" que és apte pels infants i robust davant artefactes de moviment en el context de la interacció ambulatòria de de cos sencer. Amb l'objectiu d'integrar comportaments manifestos observats amb altres fonts de dades, hem dissenyat i desenvolupat un nou protocol de codificació de vídeo i una graella de codificació adaptada concebuda per Comportaments d'Interacció Social (CIS) en infants TEA. Utilitzant un disseny de mesures múltiples, s'han recollit dades de setanta dos infants (36 díades TEA / no-TEA) de la ciutat de Barcelona, amb edats d'entre 8 i 12 anys (N=12 dones, N=60 homes). Les dades d'aquests assajos s'han organitzat en una base de dades pública i s'han processat sobre la base d'un "pipeline" de software semiautomàtic desenvolupat dins d'aquest projecte. Basant-nos en aquestes dades, hem desenvolupat tres models computacionals diferents per modelar CIS en infants TEA durant les sessions de Lands of Fog, comparant-les amb les sessions de LEGO. Els resultats d'aquesta recerca recolzen la idea de que els entorns de RM d'interacció de cos sencer són capassos de fomentar els CIS en infants TEA amb un èxit similar al de LEGO, amb l'avantatge addicional de ser més flexibles. Les troballes descrites aquí donen nova llum sobre el desenvolupament d'una eina que mediatitza, guia i recolza el progrés dels infants en la pràctica de CIS i aporta estructura i ajut als terapèutes.

Özet

Günümüze kadar yapılmış olan akademik çalışmaların gösterdiği üzere, vucut merkezli interaktif deneyimlerin tasarlandığı Karma Gerçeklik (KG) sistemleri, Otizm Spektrum Bozukluğu (OSB) olan çocukların sosyal etkileşiminde önemli bir rol oynamaya başlamaktadır. Fakat, bu sistemlerden toplanan veriler temel olarak çok modlu ve analizi kompleks olmaktadır. Bu tipdeki verilerin, veri kaynaştırması ve analizi işlemleri aracılığıyla incelenmesi, etkileşimli anlam çıkarma adına çok önemli bir adım teşkil etmektedir. Bu doktora tezinde, Vucut Merkezli Etkileşim Laboratuvarında (FuBIntLab) geliştirilmiş olan Lands of Fog deneyiminin etkinliğini anlamak adına, vucut merkezli interaktif etkileşim karakteristikleri göz önünde bulundurularak, yeni çok boyutlu veri toplama ve analizi teknikleri geliştirilmiştir. Bu büyük ölçekli karma gerçeklik ve vucut merkezli etkileşim ortamı, çocukların birbirleriyle yüz yüze oyunlar oynayabileceği ve aynı zamanda ortamda bulunan fiziksel ve sanal gerçeklik elementlerini keşfedebilecekleri bir deneyim sunmaktadır. Bu tez çalışmamızda, spesifik olarak, Lands of Fog ile Kontol Değişkeni olarak seçilen LEGO yapı oyuncağı ortamını karşılaştırmaya olanak sağlayacak bir deney düzeneği geliştirilmitir. Bu deney düzeneği, psikofizyoloji tabanlı verilerin, gözlemlenen açık davranış biçimleri ve oyun aktivitelerini içeren sistem kayıtları ile eş zamanlı olarak toplanmasına olanak sağlamaktadır. Psikofizyolojik kaynaklardan gelen verileri doğru bir şekilde toplamak adına, çocuklara uygun ve hareket kaynaklı sinyal gürültüsüne dayanıklı giyilebilir bir teknoloji geliştirdik. Gözlemlenen açık davranış biçimlerini diğer veri kaynakları ile entegre etmek için, sosyal etkileşim davranışlarının göz önünde bulundurulabildiği, yeni bir video kodlama protokolü geliştirdik ve bir video kodlama cetvelini çalışmamıza uyarladık. Tekrarlı Ölçüm Deseni kullanarak, Barselona şehrinde, yaşları 8-12 arasında olan (N = 12 kız, N = 60 erkek), toplam olarak yetmiş iki çocuktan (36 OSBli/OSBsiz çift) deneylerimizde veri topladık. Toplanan veriler organize edilerek açık bir veritabanında paylaşılmış olup, bu proje kapsamında geliştirilmiş olan bir yazılım aracılığıyla işlenmiştir. İşlenen bu veriler aracılığıyla, OSBli çocukların sosyal etkileşim davranışlarını Lands of Fog ve LEGO ortamında karşılaştırmalı olarak incelemeye olanak sağlayacak üç ayrı makine öğrenmesi modeli geliştirdik. Bu araştırmanın sonuçları, vücut merkezli interaktif deneyimlerin tasarlandığı Karma Gerçeklik (KG) sistemlerinin, OSBli çocukların sosyal etkileşim davranışlarını teşvik etme adına LEGO ortamı kadar başarılı olduğunu göstermekle beraber, bu sistemlerin değişik koşullara uyarlanabilirliği adına daha avantajlı bir noktada bulundugunu göstermektedir. Raporlanan sonuçlar, OSBli çocukların sosyal etkileşim becerilerini pratik yapmasını destekleyen, bu eyleme aracılık ve rehberlik eden, terapistler tarafından kullanılabilecek, yardımcı araçların geliştirilebilmesine ışık tutmaktadır.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

The characteristics that define computational knowledge about human behaviors has been of particular interest to researchers for many years in application domains such as healthcare and human-computer interaction. There are particular challenges to tackle the complexity of human behaviors, as behaviors are diverse, emotional and social, i.e. in relation to another person, object, or environment. In HCI, to understand the individual behaviors exhibited during interaction, an understanding of the dynamics of the interaction is necessary. These posed challenges are currently being addressed by interdisciplinary approach which draws knowledge from several research fields like computer science, psychology, sociology and medicine in an emerging field of research alternatively referred to Social Signal Processing [Vinciarelli et al., 2009] and Computational Behavioral Science [Rehg et al., 2013]. This latter will be the term used in this thesis.

One emerging application area for Computational Behavior Science (CBS) is characterising the potential rules of children interactive behavior in the view of Autism Spectrum Condition (ASC) [Bone et al., 2015, Bone et al., 2016]. ASC emerges gradually over the first years of life that affects 1 in 160 children [Elsabbagh et al., 2012], affecting 11.8/1000 in Spain [Chiarotti and Venerosi, 2020]. Records show that prevalence estimates have increased over time. ASC is characterized by challenges in social communication, refusal of new settings and the presence of restricted/repetitive behaviors [Association, 2013, Baron-Cohen, 2008]. Especially, those related to social-emotional reciprocity, communication and social interaction might lead them to have difficulties initiating or responding in social interactions.

Improvement of daily lives of people with ASC can benefit from well-structured interventions at a behavioral level [Zachor and Ben Itzchak, 2010]. Interventions that provide practice on specific skills may help people with ASC overcome some of the challenges that our society imposes on them. For example, interventions that help them acquire better social skills can make a huge difference in their daily well being, in their social integration, and ultimately in their autonomy as adults [Rao et al., 2008].

Social initiation is the primordial social act to improve social skills in general and lead a more autonomous life. Providing tools and intervention methods to foster and practice social interaction behaviors (SIBs) can help people with ASC in their social life [Strain et al., 1979]. Early intervention has the greatest possibilities of having an impact on their behaviors and on improving their adult lives, hence addressing children is essential [Carr, 1977, Baron-Cohen, 2008]. Typical interventions to foster SIBs mediated by a therapist are affected by subjectivity of the expert, and suffer the interference of having a human social agent in the process; i.e. experts are often unsure whether the ASC child's behavior is indeed an initiation or merely a response to a previous action of the expert [Srinivasan et al., 2016].

Computer-mediated intervention has the advantage of reducing bias, subjectivity, and very importantly, allows the mediation of social experiences without the human interference, as the therapist can observe the session externally [Alcorn et al., 2011]. In addition, computer-mediated interventions are stable and predictable environments which allow reducing the anxiety that people with ASC may present during real social interactions [Moore and Calvert, 2000].

Full-body Interaction experiences, based on technological configurations within the realm of Mixed Reality (MR) provide many of the advantages related to embodied cognition [Wilson, 2002]. For example, they allow users to manage physical space during their experience, fostering exploration and navigation of that space in which users share physical and virtual characteristics and objects. Users can change their point of view and adopt that of another user which allows them to compare those viewpoints and generate empathy by literally and metaphorically standing in the shoes of others. These systems also allow users to naturally apply social distance when interacting with other users and hence we can observe issues related to comfort, privacy, confidence, anxiety, etc.

1.1 Problem Statement

Full-body Interaction experiences, based on technological configurations within the realm of Mixed Reality started playing an important role in addressing the issue of encouraging socialization behaviors in children with ASC [Pares et al., 2005, Tolentino et al., 2009, Porayska-Pomsta et al., 2012, Casas et al., 2012, Antle et al., 2009, Howison et al., 2011]. However, the data from these tools is multimodal in nature and is hence complex to analyze. Fusion and analysis of this data is crucial since it offers the possibility of having a complete understanding of how these resources interact with each other.

In this PhD Thesis, given the characteristics of full-body interaction systems, we want to explore their potentialities through developing new multimodal data gathering and evaluation techniques for developing and validating assistive technologies for helping children with ASC.

Machine learning techniques have been applied for fusion of multimodal behavioral data and showing promising results in terms of extending this role [Jaimes and Sebe, 2007]. The body of knowledge already developed through modality recognizers which use machine learning already includes domains such as psychophysiology, speech recognition, facial expression analysis, gesture recognition or eye tracking [Dumas et al., 2009].

1.2 Approach

MR systems -such as the one developed in our Full-Body Interaction Lab- provide full-body, shared, face-to-face experiences in which no technological elements get in the way of the physical interaction between users, despite they might be interacting with virtual objects. This also allows users to use and detect nonverbal communication which has increasingly been understood by research as a very important part of social communication in ASC. This natural interaction that MR systems provide demands no special preparation from the users, no invasiveness of their bodies with sensors, wires, or bulky gadgets such as Head Mounted Displays (HMDs). They also do not demand special training since the children can enter the interaction area and start moving and gesticulating naturally and the external sensors will track them immediately. The children can therefore concentrate on the actual activity that is being proposed in the experience. Moreover, these experiences provide an ecologically valid context for social play as they are designed as playgrounds where the spontaneous meeting and play activities are analogous to those found in school playgrounds or urban public parks.

The FuBIntLab (Full-Body Interaction Lab) have developed in the last seven years a large-scale mixed reality (MR), full-body interaction environment, which allows two children to play face-to-face using: exploration of the physical and virtual worlds simultaneously; body gestures and nonverbal communication; joint attention; and collaborative activities. This environment allows an ASC child to play with a non-ASC peer and fosters SIB to allow the ASC child to understand the mechanisms and benefits of SIB. As described before, it is an ecologically valid context as it resembles the encounters that ASC children may find in a public park or in the school playground with a technology that is non-invasive and unencumbered. It also allows the non-ASC child to see the ASC child as a valid play partner and enhances the integration of ASC children in society.

The environment was designed in a two-years feasibility study that allowed FuBIntLab to prove that it fosters SIBs in ASC children. We then started a phase to compare its potential with a typical intervention used by therapists, based on construction toys (e.g. LEGO bricks). In this context, the goal of the present research has been to assess the design of this initial proposal by incorporating more data sources from the users in the play context to achieve a richer multimodal analysis that allows us to extract more knowledge on how our system modulates the behaviors of the children and how well it fosters SIBs. Our research has been focused on exploring and specifying the potentialities of this medium from a computational approach based on psychophysiology, body cues, video coding of overt behaviors, and self assessment questionnaires, with the use of complex statistical inference and machine learning techniques. The knowledge gained from this research has contributed to the development of effective monitoring and evaluation methods in understanding how full-body interactive systems can help ASC children improve in SIBs.

1.3 Research Objectives

To carry out this research work, an understanding was first developed of the current state of digital intervention tools in ASC. Specific focus was placed on tools which utilize virtual environments for ASC social skills intervention using full-body interaction. A research plan was then developed to better understand the impact of evaluation techniques in these interventions, specifically using multimodal evaluation and machine learning techniques.

The first phase of research included the assessment of the currently used evaluation techniques used for validating the proof of concept large scale full-body interactive environment called Lands of Fog. This environment was designed in a two-years feasibility study and it was discovered that children were in fact interacting with their peers increasingly over the course of the sessions. However, the team wanted to understand the changes in the internal state of the children during interaction to relate it to the observed overt behaviors and the reactions of the system. In this regard, psychophysiology may also be useful and the impact of anxiety might be quantified through physical cues such as heart and electrodermal activity [Hirstein et al., 2001, Kushki et al., 2013]. I joined the project in this stage and started integration of psychophysiological measures in the data collection to better understand the children's internal state while playing. At this stage, the team also proposed to develop a second version of Lands of Fog and compare its potential to that of a typical intervention used by therapists. In this regard, I designed and developed an experimental setup which allowed the comparison of two full-body interaction environments (new version of Lands of Fog and traditional therapy setting based on the LEGO construction tools) with the help of psychophysiological measures. Moreover, despite the importance of psychophysiology in ASC research, the reason for remaining paucity of evidence in this field of research on full-body interactive environments could be related to concerns around the robustness of monitoring devices. In this regard, this experimental setup included also the development of a wearable for children (both ASC and non-ASC) to capture accurate physiological data in a friendly manner, making it robust to variations and muscle activity in ambulatory contexts, such as that of full-body interaction. The addition of psychophysiological measures to the other types of data collected from users and the system, meant that the rest of data should also be time series data and therefore timestamped to allow a proper multimodal merging of all data. Because of this, I have led the development of a new method for video coding the children's overt behaviors in timestamped event coding which takes into account the range of behaviors implied in ambulatory mixed reality settings, as opposed to the usual video coding strategies developed for traditional intervention settings that only provides the frequency of events during a certain period of time. On the other hand, in order to better relate psychophysiological measures with the general state of the children during play, we adapted a questionnaire called STAIC (State-Trait Anxiety Inventory for Children) [Spielberger, 2010] to a tablet-based digital configuration and adapted a number of relaxation exercises for our setup. With this setup, I contributed to running trials and collecting data, and helped create collaborations with the ASC community and centers in Barcelona; along with the support of the experts at Hospital Sant Joan de Déu. When research is conducted in a data-driven collaborative way (especially research in ICT for ASC), it can be useful to many researchers to compare results and to build research in different directions. In this regard, I organized the data collected from these trials to be uploaded to a multimodal database following the open access policy for the MdM [1] DTIC-UPF Strategic Research Program. Considering the potential of each aforementioned modalities in helping to understand the effectiveness of full body interaction environments in fostering SIBs in children with ASC, fusion and analysis of this multimodal data could be even more important since it allows these resources to interact with each other too. In order to preprocess, classify and analyze this multimodal data, I developed code projects and end to end software pipelines. This allowed me to develop different computational models and apply machine learning techniques which allow us to understand the relation between (1) the children's internal states and anxiety levels, (2) the external behaviors shown by them (coded from video recordings), and (3) the interaction decisions that our Full Body Interaction system makes.

I also undertook a three months research stay in Cambridge (UK) at the "Affective Intelligence & Robotics Lab". Through my research stay I analysed which other data modalities could be useful to analyze and improve the computational models and provide better results for our data. We saw that the recordings of the sessions that were used for video coding could have an extended use. Therefore, I investigated other cues and modalities such as gaze and body cues. I developed models which help automatically detect, recognize and compare body cues in a multimodal manner and conclude whether the observed differences in social interaction moments between ASC and non-ASC children will be sufficient for a machine to learn to distinguish between Lands of Fog and a non-digital intervention. Moreover, this initiative also opened a dimension in our research where the role of interpersonal communication between ASC and non-ASC children (not just the ASC activity) started being investigated in terms of fostering SIBs in full-body interaction environments.

This thesis includes outcomes from these evaluation techniques related to the ability of interactive technologies to foster SIBs between children with ASC and non-ASC, and a detailed analysis of the specific interaction design concepts designed for Lands of Fog.

Finally, while developing the aforementioned structured methods for evaluation in interactive technologies where children with ASC practice social interactions, this research has also explored new interaction design techniques through sound in order to contribute to the understanding of how full-body interaction environments can be improved in terms of fostering SIBs. This research was not finally integrated to the main research branch due to a number of reasons which will be detailed in its chapter 6 Exploring Sound in Full-body Interaction. However, it opened the door to starting a new research project which is now starting in the lab.

[1]Maria de Maeztu (MdM) is an excellence research funding program from the Spanish government that has funded this stage of the present research as well as my PhD fellowship (MDM-2015-0502).

1.3.1 Research Questions

This research aimed to develop structured evaluation methods for full-body interactive technologies where children with ASC practice SIBs. Specifically, the research included implementing several data gathering methodologies and provided computational models to better understand the effectiveness of full-body interactive technologies in fostering SIBs in children with ASC. Towards this, the following research questions are investigated:

1. What evaluation methods can be effective for systems which foster SIBs in multi-user, spontaneous play based systems for children with ASC, designed using full-body interactive technology?

2. How can multimodal inputs be integrated for better multimodal representation towards fostering SIBs in children with ASC through full-body interaction environments?

3. How can we improve the understanding and adoption of SIBs in general through the design of full-body interactive systems to aid in intervention strategies for children with ASC?

4. Is it possible to develop a full-body interactive ICT system, and in particular an MR system, to affect the SIBs as much as a traditional therapy setting?

1.4 Contributions

The following research contributions are made addressing the identified research questions.

- A wearable to capture accurate psychophysiological data that is child-friendly and robust to variations in contact and muscular activity in the context of ambulatory full-body interaction.

- A structured Video Coding Method to achieve reliable timestamped coding of overt social behavior for integration in multimodal time series evaluation.

- An experimental setup for comparing an MR full-body interactive system with a control condition based on LEGO construction toys, which includes: recording physiological measures synchronized with other data sources, adapted tablet based questionnaires, and relaxation exercises.

- A computational model for estimating the social initiation behaviors of ASC children during the MR full-body interactive system and LEGO control condition based on psychophysiological measurements.

- A computational model to assess the impact of an MR full-body interactive system based on individual and interpersonal body cues in fostering social initiation in children with ASC and non-ASC, and compare it with a traditional intervention approach.

- A computational model for estimating contextualized levels of SIBs in children with ASC based on psychophysiological data, system events logs, and questionnaires in an MR full-body interactive system and a LEGO control condition.

- A multimodal database that fully records all sources of data from the interaction of children and which is publicly available, in an Open Research approach, for academic use.

- A semi-automatic software pipeline for data extraction of all the data sources.

- An exploratory study in terms of understanding the effect of sound in fostering SIBs in children with ASC through full-body interaction. (Our work associated to this project was chosen among the top 12 entries in the Student Design Competition of the 2018 CHI Conference on Human Factors in Computing Systems (CHI 2018)).

1.5 Publications

In this section we present the publications derived from this PhD Thesis. We would like to clarify that in those in which the author of this thesis is not first author, he has nonetheless contributed decisively to both the research, the development of the systems and the writing of the paper.

- Sayis, B., Pares, N., & Gunes, H. (2020, October). Bodily Expression of Social Initiation behaviors in ASC and non-ASC children: Mixed Reality vs. LEGO Game Play. In Companion Publication of the 2020 International Conference on Multimodal Interaction (pp. 140-149).

- Crowell, C., Sayis, B., Benitez, J. P., & Pares, N. (2020). Mixed Reality, Full-Body Interactive Experience to Encourage Social Initiation for Autism: Comparison with a Control Nondigital Intervention. Cyberpsychology, Behavior, and Social Networking, 23(1), 5-9.

- Sayis, B., Crowell, C., Benitez, J., Ramirez, R., & Pares, N. (2019, September). Computational modeling of psycho-physiological arousal and social initiation of children with autism in interventions through full-body interaction. In 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII) (pp. 573-579). IEEE.

- Crowell, C., Sayis, B., Bravo, A., & Paramithiotti, A. (2018, April). GenPlay: Generative Playscape. In Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (pp. 1-6).

- Sayis, B., Ramirez, R., Pares, N., (Submitted) n.d. Multimodal Context-Driven Evaluation of Social Interaction behaviors in children with Autism during Full-Body Interaction : A Comparison study between a Mixed Reality and a Nondigital Intervention, Virtual Reality Journal.

1.5.1 In Preparation

- Gali-Perez, Olga; Sayis, Batuhan; Pares, Narcis Effectiveness of a Mixed Reality system in terms of social interaction behaviors in children with and without Autism Spectrum Condition. (To be submitted to EAI PervasiveHealth 2021)
- Sayis, B., Ramirez, R., Pares, N., n.d. Multimodal Event-Driven Evaluation of Social Interaction behaviors in children with Autism during Full-Body Interaction: A Comparison study between a Mixed Reality and a Nondigital Intervention. (To be submitted to IEEE Transactions On Affective Computing, TAC)
- Sayis, B., Crowell, C., Benitez, J. P., & Pares, N. A Structured Video Coding Protocol for Reliable Timestamped Coding of Social behavior for Integration in Multimodal Evaluation: A Case in Mixed Reality for Autism. (To be submitted to Human–Computer Interaction Journal)
- Sayis, B., Ramirez, R., & Pares, N., Autism Spectrum Condition Multimodal Embodiment Open Repository (ASCMEOR). (To be submitted to 9th International Conference on Affective Computing and Intelligent Interaction (ACII))
- Sayis, B. & Pares, N. A wearable design, specifically for children with ASC, to collect reliable (resistant to movement artifacts) physiological data in full-body interaction environments (To be submitted to IEEE Pervasive Computing)

1.6 Structure of the thesis

This thesis consists of seven chapters, including this introduction. We give a brief structure of the remaining chapters of this thesis as below:

Chapter 2 - Research Context: We present a detailed literature review in this chapter. We first provide background information on ASC, challenges and intervention strategies. We review the state of the art of ICT designed for ASC. We follow with an introduction to the theory and our approach to contribute in the field of full-body interaction. After this discussion, we discuss specific evaluation techniques for developing social interventions for ASC. As a next step, we present an overview of the state of the art in the computational approaches for ASC social interventions with specific focus from the perspective of multimodal evaluation. Other topics reviewed also include data sharing in ICT based ASC interventions.

Chapter 3 - Research Framework: In this chapter we discuss the materials and methods being used for developing structured evaluation methods for full-body interactive technologies, where children with ASC practice SIBs. As a first step, we present the full-body interactive environments being tested in this thesis. We present Lands of Fog, a large-scale full-body interactive mixed reality environment, which was designed and developed previously by FuBIntLab, along with its newer version which we developed during this thesis. Next, in order to compare the potential of the new version of Lands of Fog with a typical intervention used by therapists, we present an adapted version of a commonly used intervention based on LEGO bricks that we designed in this thesis. We also discuss in detail the design and development process, physical interface, game mechanics and interaction design used for all of these environments. After this, we present the data acquisition stage used for this work with the type of data collected, definitions to this data, and the instruments that we developed and used. Finally, we explain how we collected this data in the experimental design stage

Chapter 4 - Data Analysis: In this chapter we introduce the data annotation, preprocessing and feature extraction stage for all the data sources collected in this work. This includes the details of the semi-automatic software pipeline developed for data extraction and the creation of a multimodal database from all these data sources.

Chapter 5 - User Activity Modeling: In this chapter, we present computational approaches being used for modelling SIBs in children with ASC in our MR environment compared to the traditional therapy setting based on psychophysiological, body cues, self assessment and game activity measurements. We present and discuss the results step by step. These discussions include the ideas presented in the previous chapters, linking concepts from the perspectives of design, implementation, and evaluation of the full-body interactive MR environment for fostering SIBs in children with ASC.

Chapter 6 - Exploring sound in Full-body interaction:. In this chapter, we give information on the methodologies used to explore new interaction design techniques through sound in order to contribute to the understanding of how full-body interaction environments can be improved in terms of fostering SIBs. We present an exploratory study and describe the design process, the development and the evaluation of a full-body interactive soundspace. With this additional research, we demonstrate that full-body interaction is a suitable approximation for developing systems to engage users and motivate them to create SIBs with the help of sound.

Chapter 7 - Conclusions: In this chapter, we give a summary of the research. We comment on the main contributions of the work and the limitations and the future directions of the current research.

Chapter 2

RESEARCH CONTEXT

2.1 Introduction to this chapter

We open this chapter with a brief summary of our Ethical Approach. As a next step, we introduce Autism Spectrum Condition (ASC) and the particular characteristics that affect the lives of these individuals. Since the work in this thesis is aiming to better understand the effectiveness of a medium that focuses on social interaction, we concentrate primarily on the challenges encountered in social environments by individuals with ASC and early traditional intervention strategies addressing these challenges. Next, in the context of motivating social interaction, we review the state of the art of ICT designed (non full-body interaction approaches) for ASC. The following section is concerned with Full-body Interaction experiences, based on technological configurations within the realm of MR which is the groundwork of this thesis. We address the potential of this HCI paradigm in working with children with ASC, while laying out the theoretical dimensions of embodied cognition. We look at how these projects fit into the broader range of ICT for ASC interventions and review associated work that inspires the development of the interaction environment assessed in the chapters that follow. Next we introduce subjective (e.g video coding and self-reports) and objective evaluation strategies (e.g psychophysiology and computer vision analysis) in ASC and their potential for evaluating full-body experiences. Along with this, we present an overview of the state of the art in the computational modeling and machine learning approaches for ASC social interventions with specific focus from the perspective of multimodal evaluation. Other topics reviewed also include data sharing in ICT based ASC interventions.

2.2 Ethical Approach

All our research is within the EU Responsible Research and Innovation policy (RRI) (https://ec.europa.eu/research/innovation-union). RRI is defined as "a higher-level responsibility that aims to shape, maintain, develop, coordinate and, align existing and novel research and innovation-related processes, actors and responsibilities with a view to ensuring desirable and acceptable research outcomes."[Stahl, 2013]. One of the core requirements of RRI is privacy which is defined by The European privacy directive (95/46/EC), subsequent national legislation and the current review of this entire system with a view of developing a unified European General Data Protection Regulation. In this regard. all the data used from children in our experimental trials has followed the protocols defined by applicable EU law and ethical standards through the resources provided by UPF. Each participant's ID related information and other information types (psychophysiological data, system data, behavioral data) is associated with a unique alphanumeric ID. To protect the participant's privacy, the alphanumeric ID file which is associated with participants' IDs are stored separately from the rest of the documentation in a password-protected disk in UPF. Only the researchers of the project has an access them, under the supervision of the principal researcher. The details of the ethical procedures followed can be found in section 3.4.2 Ethical Approval.

2.3 Autism Spectrum Condition

Autism Spectrum Condition (ASC) is a neurodevelopmental condition which is characterized by deficits in social communication abilities and repetitive and restricted behavior patterns [Association, 2013]. The specific signs of ASC can be observed during the first years of life and be reliably diagnosed by 2 years of age. Although DSM V and other standard references to Autism use the term Autism Spectrum Disorder (especially in the USA), the term "disorder" is increasingly understood as demeaning, especially by the people with autism themselves. The notion that they have a disorder puts them in a position of having to be cured from it. Many people with autism are defending the position that they have a condition which is only one from the myriad human conditions that exist; e.g. some people might be dyslexic, some might be hyperactive, some may have difficulties concentrating, etc. In other words, they do not like to be placed in opposition to a "typical" (or "neurotypical") person because they claim there is no actual typicality. This position is increasingly being picked up in the field and we completely agree with it. Therefore, in this thesis we will always use the term Autism Spectrum Condition and "people with ASC", as well as, "people without ASC" or "non-ASC people" to describe those that have not been diagnosed as ASC.

One of the most dominant features common to individuals with ASC are the social and communication difficulties that our society imposes on them [Ploog et al., 2013]. Even if people with high-functioning ASC, or individuals who exhibit common ASC indicators and have an IQ of 70 or greater, perform equally well in structured social tests compared to non-ASC individuals with similar IQ and age [van der Geest et al., 2002], they do perform differently in more spontaneous natural social settings [Klin et al., 2005].

Social interaction behaviors (SIBs) include person to person contact in the form of a changing sequence of initiations and responses between individuals or groups [Goffman, 1955]. A considerable amount of literature has been published on high functioning children with ASC and their capability to respond and maintain a social interaction when an initiation is addressed to them. However, they are significantly less successful in generating social initiations by themselves [Sigman et al., 1999] which can lead to reduced social interaction [Nikopoulos and Keenan, 2003]. Social initiation is the primordial social act to improve social skills by adding the possibility of starting this interaction and leading a more autonomous life. Moreover, it has previously been observed that when the frequency of social initiations increases, it is possible to see improvements in social behaviors [Strain et al., 1979].

Although anxiety is not a recognized diagnostic criteria of ASC, it is reasonable to take into account that social difficulties experienced by individuals with ASC could lead to anxiety [White et al., 2009]. If a child has a related anxiety disorder, the overall social deficiency associated with ASC may be exacerbated. For instance, social anxiety can lead to avoiding social environments, interaction with peers, and cause alienation from peers of the same age [Myles et al., 2001]. Therefore, with ASC's social difficulties, anxiety might have bidirectional relationships. Recent evidence suggests that there are relationships between anxiety, sensory hypersensitivity [Pfeiffer et al., 2005], and level of social impairment [Bellini, 2006].

Current improvement of daily lives of people with ASC can benefit from well-structured interventions at a behavioral level. Early educational interventions have been shown to yield benefits in ASC children's future health and well-being by working on challenging behaviors [Carr, 1977] and encouraging the learning process of social skills [Baron-Cohen, 2008]. In particular, the existing body of research on early intervention in children with high functioning ASC suggests that reports outlining the progress of children have a positive trend [Zachor and Ben Itzchak, 2010, Rao et al., 2008]. Play therapy is one of the most widely used intervention strategy for the improvement of social and communication skills [Casenhiser et al., 2013, Green et al., 2010, Kasari et al., 2008, Mahoney and Perales, 2003, Pajareya and Nopmaneejumruslers, 2011]. However, it should be noted that in imaginative and symbolic play with toys, children with ASC appear to exhibit unusual behaviors. They may play with toys in a repetitive manner and get completely self-absorbed. Their approach to objects involves exploratory and manipulative behavior while being more attached to these objects compared to their non-ASC peers [Rowland and Schweigert, 2009, Williams, 2003]. Yet play is still an important developmental tool for children with ASC as it is for children with non-ASC, but different strategies might be taken into account in play therapy. One of the best known and most promising studies on play-based therapy for children with ASC was conducted by Legoff in which he used LEGO bricks [LeGoff, 2004]. Each child played a particular role in the construction process in experiments, which required all children in each group to work together to achieve the final objective. Their findings showed improved acquisition of social skills, particularly when peer interaction was initiated. One longitudinal study also obtained promising results in improving the acquisition of social skills in children with ASC compared to children without ASC [Legoff and Sherman, 2006]. The other most popular intervention to enhancing individuals with ASC's daily well-being is the Applied behavioral Analysis [Cebula, 2012, Johnston et al., 2006, Estes et al., 2014]. There is a large volume of published studies describing the role of ABA techniques in fostering social behaviors, play skills, and communication abilities [Dawson et al., 2012, Eapen et al., 2013, Grindle et al., 2012]. ABA is based on teaching skills and modifying behavior using a reward system that personalizes every child's most meaningful reward to replace undesirable behaviors. However, over the years since ABA has been criticised because of these rewards potentially causing conditioning issues, ABA has started being applied in a more flexible way so that children can generalize their learning to other contexts [Devita-Raeburn, 2016].These reinforcements may not only take the shape of verbal feedback but also include pleasant visual and audio cues. In this regard, information communication technologies (ICT) might have a unique potential due to its interactive nature where immediate audiovisual feedback is possible whenever user activity comes along with therapists' expectations.

2.4 ICT for Autism

It has been demonstrated that computer-based tasks are one of the activities that people with ASC may have a tendency to enjoy [Chen and Bernard-Opitz, 1993]. In an investigation into improving children with ASC's desire to engage, Brown and Murray [Brown and Murray, 2001] reported that children with ASC have a natural affinity for ICT probably due to its stable and predictable aspects. During the past 20 years, the emergence of technology and promising results on the affinity that children with ASC have towards ICT has allowed the development of more engaging and dynamic interventions and learning experiences. There is a large volume of published studies describing the important role of ICT applications in ASC therapy because it triggers less socially threatening situations [Brown and Murray, 2001, Bernard-Opitz et al., 2001, Moore and Calvert, 2000, Hourcade et al., 2013]. Moreover, it has been seen that computer-mediated intervention has the advantage of reducing bias, subjectivity, and very importantly, allows the mediation of social experiences without the human interference, as the therapist can observe the session externally [Ramdoss et al., 2012]. Typical interventions to foster SIBs mediated by a therapist are affected by subjectivity of the expert, and suffer the interference of having a human social agent in the process; i.e. experts are often unsure whether the ASC child's behavior is indeed an initiation or merely a response to a previous action of the expert [Srinivasan et al., 2016]. ICT techniques are therefore considered suitable to support typical therapy environments, as change can easily be regulated [Alcorn et al., 2011] and feedback can be programmed to be on the spot and stable [Moore and Calvert, 2000].

Much of the current literature pays particular attention to social robotics which mainly focuses on motivating socialization through imitation and joint attention. Various studies have assessed the important impact of imitation in development of social communication skills [Ingersoll, 2008, Scassellati, 2005]. Moreover, robots can make social demands more predictable, thus avoidance towards unfamiliar or unpredictable social situations which are commonly seen among individuals with ASC might be reduced. It has been shown that robots elicit positive responses from children with ASC [Dautenhahn et al., 2002]. The findings revealed that the children exposed to a robot mediator were more closely involved in interaction and displayed more eye contact than those exposed to a human mediator [Duquette et al., 2008]. It has been demonstrated that robots increase the joint attention between children and adults when used as mediators [Robins et al., 2004]. However, it should be noted that robots being artificial entities might not provide the ecological validity desired [Scassellati et al., 2012], in contrast to a face-to-face play environment between ASC and non-ASC children, such as ours that provides strong ecological validity.

Research on Tangible User Interfaces (TUI) for intervention for ASC children is also seen as an area of great promise as the results from several studies obtained show positive outcomes when TUI is used as a mediator for fostering learning of communication skills [Marwecki et al., 2013, Villafuerte et al., 2012, Farr et al., 2010, Ortega et al., 2015]. Results from these studies suggest that the use of mediating objects makes "computational offloading" for cognitive processes simpler, as a result creating a space for children to understand the actions and intentions in social interaction easier. The visual and tactile properties of tangibles may help focus children's attention on social interaction and allow children to create a common context in terms of sharing ideas and objects.

The potential of Virtual Environments (VEs) has been also widely investigated for teaching social skills to children with ASC [Parsons and Mitchell, 2002, Kerr et al., 2002, Schuller et al., 2013]. Regardless of the physical interface, VEs can be used as adequate intervention environments due to the flexibility, adaptivity and intelligence provided. The users can practice skills safely, and explore situations and interact with virtual objects in scenarios that can be similar to real life scenarios. In these environments the stimuli the user receives can be fully controlled thus creating an ideal mediation for fostering communication and learning for ASC children. In an analysis of VEs, Mitchell et al investigated different VEs in terms of their potential in unfolding social understanding in children with ASC and found improvement in their social skills independent from the virtual contexts provided [Mitchell et al., 2007]. In the VE research, it is also possible to see that virtual agents can form the central focus to be used to teach social competence and communications skills to children with ASC [Gray and Garand, 1993, Hopkins et al., 2011]. Similar to the research on interactive robots as social partners, digital peers can help ASC children to have a better understanding in social settings.

Virtual Environments can be delivered in different levels within the Reality-Virtuality (R-V) continuum [Hopkins et al., 2011]. For example, Virtuality and Reality can be positioned at opposite ends of the RV Continuum which relates purely virtual environments to purely real environments. In an investigation into Virtual Reality and its potential effect in fostering social interaction in high functioning ASC, Fengfeng and Tami [Ke and Im, 2013] used activities that involved recognizing gestures and facial expressions of avatars interacting with them in a school cafeteria and communicating with them during a birthday party. They obtained positive results in terms of increase in social initiations, responses and greetings during the intervention and observed improved social competence measures after the intervention. On the other hand, although VR approaches are promising, some aspects of full immersion may play a substantial role in avoiding VR, particularly for children who are more sensitive to immersive VR exposure compared to adults. For example, in some cases invasiveness is not desired because some children might reject using immersive VR delievered through HMDs, which would be quite problematic in case they suffer a meltdown. Moreover, such an immersion actually disembodies them from the body they are getting to know during their development. This hinders their propioception, their possibilities of using their bodies in a natural way, and therefore does not allow them to use non-verbal communication cues. They may also suffer VR nausea effects and/or have motor issues while interacting with the HMD on. The alternative of using Augmented Reality (AR) through non-immersive devices could deem an adequate alternative. [Feiner et al., 1993, Azuma, 1997, Buchmann et al., 2004, Bai et al., 2013]. However, AR based on mobile devices such as tablets or smartphones keep the hands of the users busy and the focus on the screens become actually digital dividers [Betsworth et al., 2014] that separates the children from the physical world. In other words, the children would not be able to have a truly face-to-face interaction with their peers.

It should be noted that, much of the current AR-VR research has been on experiences delivered through handheld/heads-down/HDM devices. In these the children are either literally separated from the physical world and the other users, or at least their focus of attention is placed on the device. Although they can connect with remote users, the type of communication and interaction between the users mostly do not incorporate full-body interaction through which nonverbal communication can also be possible.

2.5 Full Body Interaction and Mixed Reality

As technological progress has permitted the spread of interactive technology beyond the conventional mouse and keyboard interfaces, users can benefit from more natural and intuitive experiences and are not restricted to tethered or invasive physical interface elements [Chen, 2012, Tajadura-Jiménez et al., 2017]. In particular in the RV framework, it is possible to define a generic Mixed Reality (MR) environment as one in which real world and virtual world objects are presented together. Our approach to this is based on spaces that use: large-scale projections (on the floor and/or special screen walls); computer vision tracking of users and/or objects they may manipulate; spatialized sound; etc. With such MR systems, it is possible to provide full-body, shared, face-to-face experiences in which no technological elements get in the way of the physical interaction between users, despite them being able to interact with virtual objects. Users may also navigate physical and virtual space simultaneously exploring and discovering elements, changing their points of view and using their bodies to understand the scale of the world they are experiencing. This also allows users to use and detect nonverbal communication which has increasingly been understood by research as a very important part of social communication in ASC. This natural interaction found in MR based full-body interaction systems demands no special preparation from the users, no invasiveness of their bodies with sensors, wires, or bulky gadgets such as headsets (e.g. in VR or AR).

The potential of MR based full-body interaction systems is based on cognitive theories that consider physical interaction and cognitive processes working together, called Embodied Cognition. Embodied cognition theories emphasise the formative role of proprioceptive and kinesthetic cues when we make sense of the world through our bodies [Wilson, 2002]. According to the Embodied Cognition, human awareness is based on the active interplay between our body and the environment [Borghi and Cimatti, 2010]. Along with the unique viewpoint, each individual's cognition is affected and linked with its respective body dynamics and social context [Roussos et al., 1999]. Activity theory (deriving from Vygotsky's social contructivism [Vygotsky, 2012]) suggests that intra-psychological mechanisms arise in a relevant environmental context during activity between people and objects (inter-psychological processes) [Wertsch, 1983]. Sense of spatial organization and navigation through physical environments provide a greater comprehension of one's surroundings than simply viewing a two-dimensional representation [Bartoli et al., 2013]. On the other hand, while there is no consensus on the question around "Do people with ASC have distinct embodiment?", there are several studies examining this matter [Donnellan et al., 2013, De Jaegher, 2013]. There is a great deal of interest in the various ways of moving, perceiving, and emoting of people with ASC [Gepner et al., 1995, Gepner et al., 2001, Baranek, 2002, Rogers and Ozonoff, 2005, Fournier et al., 2010, Whyatt and Craig, 2012].

Taking the potential of ICT in the context of interactive computer controlled intelligent systems for ASC children and the benefits of Mixed Reality settings as naturalistic and ecologically valid worlds in which therapy can be done, we can argue that MR based full-body experiences as a blend of systems that take advantage of situating the body in virtual world and the physical world on the same hierarchical relationship that they can coexist, may indicate a greater potential for intervention. Much of the current literature on full-body interaction intervention for ASC children pays particular attention to the acquisition of agency, the learning of specific skills, the mediation of communication between therapist and patient, and the learning of social communication skills [Pares et al., 2005, Tolentino et al., 2009, Porayska-Pomsta et al., 2012, Casas et al., 2012, Antle et al., 2009, Howison et al., 2011].

Already in 2001-2004 a pioneering project in multimodal interactive spaces for children with ASC, known as MEDIATE, integrated the concepts of full-body interaction where children were able to involve themselves within the environment through body cues, touch and movement [Pares et al., 2005]. In the ASC context, this study was one of the first to discuss ways to encourage and improve exploration in Mixed Reality settings in both communication and social scenarios. In 2009, The project SIIMTA (Real-Time Full-Body Interaction System and Music Therapy for people with disabilities or disorders such as Autism in Catalan) was developed to investigate the key aspects of full-body interaction technologies with music therapy for children with ASC [Mora Guiard, 2009]. The Pictogram Room was another full-body interaction videogame to help ASC children recognize body posture and gestures in which an augmented digital mirror, based on a Kinect sensor and a display in front of the user was designed. Users could see themselves together with a caregiver in the display where the system adds a number of digital elements (such as stick body diagrams) to the body gesticulation of children to easily communicate the relation between diagrams and body [Casas et al., 2012]. Another full-body interactive project for acquiring social skills was the ECHOES project [Bernardini et al., 2014]. In this study, researchers described how the actions of a child were continuously interpreted by a system based on artificial vision that tailored a virtual charachter's responses to interact with the child spontaneously. The system thus made it possible to conduct tasks where joint attention, such as pointing, could be exercised. Keay-Bright et al. [Keay-Bright, 2007] designed a system called ReacTickles for children with ASC. It was a playful exploratory environment where children could quickly discover various "magical" experiences without previous knowledge of the technology. The goal of ReacTickles was to explore the potential of virtual environments to encourage expression and to facilitate immersion during playful learning intervention. Another initiative aimed at helping ASC children learn social skills was Autinect [Agarwal et al., 2012]. The project questioned the potential of a gesture-controlled virtual peer in teaching social skills. Bhattacharya et al. [Bhattacharya et al., 2015] aimed to encourage engagement with peers and social interaction in children with ASC in a classroom setting using Kinect. In Pico's Adventure, children with ASC were called upon to cooperate with others by synchronized gestures and motions, where game play was controlled by Kinect [Malinverni et al., 2014]. Collaborative play could be scaffolded in children with ASC corresponding to the presence of clear, structured interaction dynamics. However, although none of these studies used an HDM, the interaction design approach in most of the these setups is based on third-person interaction paradigm, where the tracking camera is placed on the same side as the visual display and hence the users need to be placed at a distance from the display+camera block to allow for a good tracking of the user. This then forces the system to represent the user within the system (with a silhouette or digital mirror image) to provide a clear feedback of where the user is actually impacting the virtual world. Hence, this causes compromised embodied interaction and face-to-face interaction between users [Parés and Altimira, 2013]. In this regard, Mora-Guiard et al. [Mora-Guiard et al., 2016] developed a large scale full-body interaction system projected onto the floor which provides shared, face-to-face experiences. In such setups, interaction is not limited to just verbal exchanges as they also include nonverbal body cues such as interactional synchrony, joint attention, and proximity. Such setups can foster exploration and navigation of that space which shares physical and virtual characteristics and objects. Users can change their point of view and adopt that of another user which allows them to compare those viewpoints and generate empathy by literally and metaphorically standing in the shoes of others. These systems also allow users to naturally apply social distance when interacting with other users and hence it is possible to observe issues related to comfort, privacy, confidence and anxiety. Practicing socialization in such an environment catered towards individuals with ASC can be a way to understand social interaction better while simultaneously forming al patterns.

2.6 Evaluation methods in ASC

2.6.1 Video coding

Manual coding of live or video-taped observations is accepted as one of the most widely used methods for testing the effectiveness of the social interventions in ASC. Typical techniques of video coding offer the possibility to do posterior analysis of captured behavioral observations from the intervention sessions. Coding can follow a lengthy process through analysis of qualitative data using an inductive (bottom-up) approach which employs the method of observation and denotation of target behaviors, which are later classified based on affinities to construct meaning. However, the most common way to do coding is to employ a deductive (top-down) approach which is relatively a more objective strategy, where independent coders watch intervention videos and mark target behavioral cues using a predefined coding scheme.

Coding schemes can be positioned on a continuum depending on the degree of inference needed [Alexander et al., 1995]. This continuum ranges from global to event-based systems, namely macrosystems to microanalytic systems [Adams, 2013]. Global assessments provide a summary judgment of behavior on specific dimensions and requires larger coding units and a higher level of inference [Malik and Lindahl, 2004]. Through splitting data into relatively short time units, microanalytic coding allows recording moment-to-moment behaviors [Tardif et al., 1995]. These defined short time units may change based on the temporal dynamics of the behavior under investigation and they can be used in different strategies ranging from interval based coding (state event) to ones which record every instance of behavior (point event). Moreover, due to its certain degree of objectivity and minimum demand for judgment, microanalytic coding is proposed to be superior to macroanalytic coding [Alexander et al., 1995, Carlson and Grotevant, 1987]. However, in microanalytic coding, because of the detailed coding requirements, multiple coders have to work together to achieve high inter-rater or inter-coder agreement and it is a resource intensive and costly process which is restricting its use to research where considerable funding is available [Morawska et al., 2015]. Moreover, in reaching inter-rater agreement scores in microanalytic coding strategies, it is common to see the usage of frequency based validation approaches where different coders only compare the similarity of the number of events each coded. It should be noted that such strategies does not fully validate the effectiveness of the developed microanalytic coding schemes since sequentiality of the events are not taken into account in the validation process of these schemes.

In order to facilitate social understanding and interaction in high functioning children with ASC, Nirit Bauminger [Bauminger, 2002] created a scheme to observe social interactions in which social initiations and responses are marked into three categories: positive social interaction, negative social interaction, and low-level interaction. The scheme was used to code behaviors during playground pre and post recess sessions in the intervention. This coding scheme was used in conjunction with other evaluation methods such as problem solving measures, an emotional inventory, and teacher reports to detect changes in emotional understanding and social skills. The scheme proposed in the Bauminger's study is based on the behavior Coding Scheme for children designed by Hauck et al [Hauck et al., 1995], which was used to investigate the frequency of social initiations in retarded children and ASC children in ecologically valid conditions. In this study, Hauck comes to a conclusion that structured play settings may provide a supportive framework for scaffolding social interactions in children with ASC. Bauminger proposes changes to fit the playground context into this framework with additional coding scheme elements such as sharing experiences/objects and social communication. The addition of social responses as a category alongside social initiations was also another significant contribution from Bauminger's research. However, although the social aspects covered in Bauminger's adaptation is well tailored from a psychological stance for interventions for children with ASC, it does not take into account environments in which an interactive system plays an active role in facilitating social interactions among children.

In the field of developmental psychology, research has historically concentrated on describing categories of play as individual progresses in social development and understanding of others. In 1932, Parten categorized social participation into six categories: unoccupied behavior, solitary play, onlooker behavior, parallel play, associative play, and cooperative play [Parten, 1932]. In 1951, Piaget specified a criteria for play and categorized play behaviors into cognitive developmental stages [Piaget, 2013]. There are also several recent observation schemes which have been developed for use in social and collaborative play environments for non-ASC children, based on a developmental psychology approach. The Play Observation Scale (POS), first separates activities into Play or Non-Play, then assigns qualifiers such as social or cognitive play which also include a sublevel definition for functional play and exploration [Rubin, 2008]. The Outdoor Play Observation Scheme (OPOS) is another coding scheme designed for evaluating children's head-up game behaviors or play behaviors comparable to outdoor games [Bakker et al., 2008]. In this coding scheme, events to be coded is differentiated into two categories namely point events and state events which take into account the temporal dynamics of the activity, while also taking into account the direction of the events (e.g looking direction). Observers watch recordings of the play sessions using this scheme to make systematic assessments of the children's physical activity, focus, and social interaction. This scheme is a step forward for analyzing the focus of children's activity; however, the social interaction parameters defined in this scheme do not explicitly address the specific behavioral characteristics of children with ASC.

In the field of full-body interaction, Autism Diagnostic Observation Schedule [Lord et al., 1999] categories have been used to categorize verbal and nonverbal behaviors presented during social game play [Bianchi-Berthouze, 2013]. Bianchi-Berthouze's codification strategy also includes additional verbal and nonverbal identifiers such as speech and emphatic gestures. Although this was not designed for ASC children specifically, and does not include attributes in the context of SIBs such as "initiation" or "response", their approach is fine grained in terms finding the differences between verbal and nonverbal behaviors, and adding an additional level of structure compared to other coding schemes, which could allow detection of nonverbal behaviors in social interventions in ASC.

When coding behaviors during interventions using an interactive system, much richer results can be achieved if the data from video coding is integrated with system data (log files of system decisions and actions) or psychophysiological data which are time series data keeping track of what is going on with the participants at every instant. This would allow the generation of a shared multimodal representation that could be used in understanding the deeper details of specific behaviors. This also implies that video coding strategies used in such interactive systems should be flexible enough to allow integration of different modalities. In this regard, since no existing observation scheme was found that was suitable and accounted for settings where the system plays a role in fostering SIBs towards or between the children with ASC, we developed an adapted observation scheme during an iterative development process. This adaptated observation scheme was based on Bauminger's study as it was already defined for interventions for children with ASC and from a psychological stance. Moreover, it has already been in use in a social skills therapy in another research project with clinical application by our collaborators from Hospital Sant Joan de Déu. (see section 4.2.1 Video Coding).

2.6.2 Psychophysiology

Psychophysiological signals can facilitate an important channel for relatedness of mind and body and complement the information provided from observable behavioral cues. Moreover, psychophysiological signals may produce useful information for people with ASC in the context of social interactions, sensory processing, and emotional expression

[Picard, 2009, Welch, 2012]. As there are some cases where people with ASC may be exhibiting overt behaviors which do not go hand in hand with their internal state, psychophysiology can help us better understand the underlying causes of such situations. Besides, psychophysiology is an important complement to self-assessment questionnaires as reports from many people with ASC alone can be misleading [Picard, 2009].

In this thesis, we are mainly focusing on the role of the Autonomic Nervous System (ANS) and its relation to psychophysiological measurements such as electrodermal activity (EDA) and electrocardiogram (ECG) signals. Thanks to their relative non-invasiveness and reliability, these signals have been widely used as markers of psychological internal states [Berntson et al., 2017, Cacioppo et al., 2007, Boucsein, 2012]. The ANS is responsible for modulating the internal and external demands in favor of keeping the body in a homeostatic state which is considered as a steady internal state. It has two divisions namely the sympathetic division (associated to fight-or-flight response) and the parasympathetic division (associated to rest and digest response) [Mendes, 2009]. The psychophysiological activity related to these divisions will be called SNS and PNS activity respectively in the rest of this thesis. In relation to ASC, these signals have been broadly studied as well [Hirstein et al., 2001, Kushki et al., 2013]; for example they have been used to understand the alteration in both Sympathetic nervous activity (SNS) and Parasympathetic nervous activity (PNS) among individuals with ASC during social interaction and it was suggested that the degree of alteration is possibly linked to the severity of social difficulties [Neuhaus et al., 2016].

Nonetheless, physiological signals are prone to contamination by motion artifacts [Boucsein, 2012]. Particular consideration should be paid, especially in full-body interaction environments, to using both ECG and EDA, since they are susceptible to motion artifacts due to movement of the peripheral body. Furthermore, an association between physical activity and the measured physiological signals can also exist [Picard and Healey, 1997]. Thus, an ambulatory nature of the experiences (especially in ecologically valid conditions) and the design of wearable sensing devices should be considered carefully in terms of noise artifacts and how physical activity of participants can affect the assessment of physiological measurements.

As explained in subsequent sections, this thesis has performed a study on the most robust approach to measure EDA and ECG for ambulatory contexts that are natural in our MR full-body interaction system. We provide: the literature review on electrode positioning for each signal; our study on the best positions given an ambulatory activity; and the design of a child-friendly wearable that allows safe and robust detection of the signals, while being attractive for the children to avoid anxiety and negative reactions towards the sensor system (see section 3.3.1 Wearable Multichannel Psychophysiology).

2.6.2.1 Electrodermal Activity

The EDA signal is made up of the superposition of two signals: the tonic level of skin conductance (SCL), representing the baseline signal, and the superimposed phasic increases in conductance. The phasic components reflect a unitary skin response (SCR). In turn, the responses are given by the activity of the eccrine sweat glands in response to external stimuli [Fowles et al., 1981]. The SCR is widely associated with autonomic markers of arousal and could be observed in social contexts [Lang, 1995, Critchley, 2002, Sequeira et al., 2009]. However, previous research findings into SCRs in ASC have been contradictory [White et al., 2014]. The evidence of hyper-arousal state in children and adolescents with ASC have been identified in several SCR studies. Kylliainen et al. [Kylliäinen and Hietanen, 2006] designed an experimental setup in which children with ASC and controls asked to view two types of gaze (directed or averted) from an adult face. Unlike the controls, in the directed gaze view, SCRs from the children with ASC were significantly stronger than the averted gaze view. In a similar study, Joseph et al. [Joseph et al., 2008] demonstrated that children with ASC exhibited significantly larger SCRs than non-ASC children to faces with direct and averted gaze. In contrast to these studies, in various studies including tasks such as face judgement [Hubert et al., 2009], viewing socially pleasant images [Mathersul et al., 2013], viewing emotional live and video faces [Riby et al., 2012], it was possible to see lower levels of SCRs in participants with ASC. On the other hand, during a risk-taking task, South et al. [South et al., 2011] compared children with ASC and non-ASC and reported no significant differences in baseline SCRs or group SCRs. Nonetheless, it was found that in the ASC group, risk taking behavior was positively associated with increased social anxiety and negatively associated with social functioning.

2.6.2.2 Cardiovascular activity

The ECG signal represents the fluctuations of cardiac potential and it is the most widely used approach to detect the variations between consecutive heartbeats, known as HRV. Measures of HRV such as Vagal tone and Respiratory sinus arrhythmia (RSA) are valuable indicators for understanding PNS responses to stress [Porges et al., 1994]. Polyvagal theory suggests that greater PNS activity of the vagal tone is associated with better social interaction and emotion regulation [Porges, 2007]. A number of studies supported this theory by showing the association between high frequency heart rate (reduced vagal control) and anxiety [Licht et al., 2009, Rottenberg, 2007, Laborde et al., 2017].

Taken together with Polyvagal Theory, there is some evidence indicating the correlation between higher PNS activity and higher social functioning in ASC [Patriquin et al., 2013, Neuhaus et al., 2014]. Moreover, lower PNS activity was associated with increased parent reported anxiety scores and lower socialization skills [Guy et al., 2014]. It has been also shown that better pragmatic language skills in children with ASC was correlated with high vagal tone [Klusek et al., 2015]. On the other hand, Levine et.al [Levine et al., 2012] did not find significant differences in vagal tone responses to a social stress test between individuals with ASC and non-ASC. Moreover, it has been also noted that the research in these matters consists only of a small body of literature and findings are diverse [White et al., 2014, Benevides and Lane, 2013].

2.6.3 Computer Vision and Body Cues

Face and body inputs provide valuable insights about behavioral cues in social interaction [Vinciarelli et al., 2009, Gunes and Schüller, 2017]. It has been suggested that children with ASC can have uncommon social behaviors such as diminished focus for social scenarios, less attraction towards gazing to faces, and suppressed emotional expression [Ozonoff et al., 2010, Flanagan et al., 2012, Brisson et al., 2012, Schuller et al., 2013, Gima et al., 2018]. In this regard, computer vision technology offers a great opportunity to understand mechanisms of ASC children's possible distinct behavior, by providing noninvasive tools for automated capture of body cues and allowing the study of interpersonal behavioral relationships between individuals [Rehg, 2011, de Belen et al., 2020].

Using Computer Vision (CV), researchers have been able to automatically quantify possible facial and body cue differences between children with ASC and non-ASC. In a recent study by Dawson et al. [Dawson et al., 2018], head pose angles (estimated through facial landmarks) were found to distinguish between children with ASC and non-ASC (with a higher rate of head movements in ASC) while they were exposed to a series of social and non-social stimuli. Using a similar social content, Martin et al. [Martin et al., 2018] evaluated the head movement differences in children with ASC and non-ASC and found distinguishable differences in yaw and roll indicators of the head pose. In another study, Ardalan, A. [Ardalan et al., 2019] investigated how whole-body movements would differ in children with ASC and non-ASC, and which body parts might drive this effect. They found that ASC participants exhibited more variability in their movements in general and no single region of the body seemed to specifically drive this variability.

When it comes to automatically analyzing nonverbal behaviors in engagement and social interaction, Sanghvi [Sanghvi et al., 2011] used upper body silhouette features to train a set of classifiers for engagement prediction in children playing with a robot. Using a similar approach, researchers have been able to predict engagement based on the children's facial cues [Anzalone et al., 2015] and body movements [Colton et al., 2009]. Esteban [Esteban et al., 2017] used gaze direction, facial expressions and body posture to classify stereotypical behaviors and social engagement of children with ASC. Rudovic et al.[Rudovic et al., 2018] achieved personalized emotional perception and engagement using contextual and personal information of children with ASC during robot-assisted therapy. Mataric demonstrated that CV can be used during free play environments to research the behavior of ASC children towards robots [Feil-Seifer and Mataric, 2010]. However, the focus of these studies was on child-robot interaction using children's behavioral cues.

In the context of social interaction between individuals, Piana [Piana et al., 2013] presented an emotion recognition system that used full-body movement features to help children with ASC. Specifically, they explored a collection of body features that can be extracted from video sequences to recognize emotions automatically. Their framework monitors ASC children while they interact with others or play a serious game and evaluates their ability to express and understand emotions. However, the dataset used to test the feature extraction algorithms was based on non-ASC individuals and was therefore not yet aligned with the final objective of helping ASC children. Rehg et al. [Rehg et al., 2013] provided a detailed study on the analysis of children's social and communicative behaviors.However, the interaction data collected in [Rehg et al., 2013] are the dyadic social interaction between adults and children with ASC. Winoto proposed an unobtrusive sensing system to observe social interactions among individuals with ASC in the classroom setting [Winoto et al., 2016]. However, the feature set they used as indicators of interpersonal relationships was relatively small and they faced technical problems with the noise of the testing environment. The project by Coppola [Coppola et al., 2016] introduced one of the most informative sets of features based on two interacting participants. Although this was not designed for ASC specifically, their analysis provides a robust automation of body cue recognition. Their results show that their approach has a potential in ASC research. In this regard, to analyze the body cues, we utilized and adapted the methodology of Coppola [Coppola et al., 2016] and extracted a set of individual and interpersonal spatio-temporal features from both subjects, using the skeleton joints. Further detail on how these have been used may be found in section 4.2.5 Body Cues

2.6.4 Self-reports

Data from several studies suggest that self-reflection skills might start appearing among children at the age of 4 which is acceptable and becomes consistent over the years [Sturgess and Ziviani, 1996, St-Laurent-Gagnon et al., 1999, Sturgess et al., 2002]. The list of aspects that can be self-reported by a child, such as emotions, health-related thoughts, discomfort, physical activity level, is still in the early stage of refinement; however, there are already a variety of methodologies for collecting child self-report [Sturgess et al., 2002].

The challenges of obtaining reliable self-reports from children with respect to their interest in interactive environments were also addressed in previous design studies. One study by Read [Read et al., 2002] presented a toolkit called Fun Toolkit for measuring fun with children. This toolkit was planned to be used for assessing the usability of a novel interface and it includes a set of tools namely Again-Again Table (measuring engagement), The Smileyometer and Funometer (measuring comparative fun), and Fun Sorter (measuring preference of activity). In another study, Smileyometer was tested in detail in the context of measuring fun in Museum, and it was found that children have a tendency to produce biased results by choosing the higher fun ratings while it was possibly not the case, shedding light into reliability of such self-report measurements in children [Sluis et al., 2008].

Obtaining reliable self-reports from non-ASC children might be challenging yet obtaining these reports from ASC children can be even more challenging as children with ASC might have difficulties in interpreting their own experiences and proprioception [Blanche et al., 2012]. There is a relatively small body of literature that is concerned with developing reliable self-report methods involving children with ASC. However, in a study done by Kirby et al., researchers emphasized the potential of using contextual information and sensory descriptors in detailing the interview questions [Kirby et al., 2015].

In all the studies reviewed, we discovered that questionnaires and interviews are still one of the commonly used approaches for obtaining feedback from ICT based social interventions in children with ASC. However, we also realized that what stands out as a trend in many projects was the usage of non-standardized self-reports to obtain information. They were customized according to the specific needs of the experimental designs, not allowing a just comparison between data obtained from different systems. However, since we are using a multimodal approach in which we are triangulating a number of different data sources, given the success of sensory objects in interviewing children with ASC in previous studies[Kirby et al., 2015], and the potential of multi-touch apps [Hourcade et al., 2013], we have included self-reports from our participants through an especially designed questionnaire based on a tablet interface. Later in section 3.3.4 Questionnaires, we explain how we have adapted some questionnaires to a format that is easy for the child to use.

2.7 Multimodal evaluation & computational modeling

In recent years there has been an increasing amount of literature on multimodal data capturing strategies. In this regard, fusion and analysis of this data started playing an important role in research since it offers the possibility of having a complete understanding of how these resources interact with each other [Streeck et al., 2011]. Machine learning techniques have been applied for fusion of multimodal behavioral data and showing promising results in terms of extending this role [Jaimes and Sebe, 2007]. The body of knowledge already developed through modality recognizers which use machine learning includes domains such as psychophysiology, speech recognition, facial expression analysis, gesture recognition or eye tracking [Dumas et al., 2009].

There has also been growing interest in recent years in integrating several behavioral modalities to produce superior results and even outperform previous state-of-the-art approaches that use only a single modality to understand ASC behavioral characteristics [Messinger et al., 2015, Bone et al., 2016, de Belen et al., 2020]. For example, the project Virtual Environment System for Social Interaction (VESSI) [Welch et al., 2010] undertook an evaluation which brought together the ratings from a clinical observer and participants physiological signals to undercover patterns in anxiety with respect to specific social factors (e.g., eye contacts and proximity). Their observations were aligned with the previous social anxiety studies in the context of non-ASC adults in real-world experiences, as well as ASC and non-ASC children in a virtual environment. Esteban [Esteban et al., 2017] used gaze direction, facial expressions, body posture and voice to classify stereotypical behaviors and social engagement of children with ASC. In another study, two computer-based cognitive tasks were used to investigate anxiety in children with ASC [Liu et al., 2008]. Using ML techniques they were able to classify anxiety (with the tagged instances from a behavioral therapist) at an average accuracy of 79.5%, using different physiological modalities together, including ECG, EDA, EMG, blood volume pulse (BVP), temperature, bio impedance, and heart sound. Rudovic et al. achieved personalized emotional perception and engagement using contextual and personal information of children with ASC during robot-assisted therapy. The type of data fused was from three main data sources (i) video recordings of body cues including facial expressions, head movements, pose, and gestures; (ii) audio recordings; and (iii) psychophysiological measurements including heart rate , EDA, and body temperature [Rudovic et al., 2018]. Javed et al. presented a computational modeling approach to quantify the level of social engagement derived from behavioral data collected from children with ASC and non-ASC. The study was conducted within a child-robot interaction setting which is based on a sensory maze game. The behavioral data collected from the subjects included video, audio and motion-tracking data and later this data used for generating personalized computational models of social engagement [Javed et al., 2020]. The focus of these studies was on child-robot interaction using ASC children's behavioral cues. Similar multimodal and computational evaluation approaches can also be applied to full-body interaction environments for children with ASC.

2.8 Multimodal data sharing in ASC research

Several lines of evidence show that ICT made it possible to develop systems that allow for a wide range of multimodal interaction scenarios and provide repeatable, subtle and naturalistic stimulus while quantitatively monitoring the performance of the children with ASC to form a comprehensive understanding of current practices in intervention approaches [Scassellati, 2005]. Considering all of this evidence, ICT studies for ASC do not appear motivated to share their data.

When research is conducted in a data-driven collaborative way, it can be useful to many researchers to compare results and to build research in different directions. Especially in ASC research, it is important to run a study in a data-driven collaborative way, since it is generally a challenge to access participants in experiments and some segments of the research requires to recruit participants from across the spectrum [Haas et al., 2016].

As important research is being undertaken in genetics to try to understand origin, cause and typification of ASC, there has been a large volume of work during the last decade for creating reusable multimodal corpus consisting in metabolomic, and genomic datasets from individuals with ASC (iHART platform, The MSSNG project, NDAR, AutDB, The SFARI Gene database). On the other hand, the available ICT tools including embodied Interaction and psychophysiology are increasingly showing potential in ASC research and data from these tools is multimodal in nature and is hence complex to store and analyze.

An example of how multimodal datasets can be gathered to model and interpret the actions of children with cognitive disabilities is presented by researchers at the Georgia Institute of Technology [Rehg et al., 2014]. In this research, the children with ASC participated in semi-structured play interactions while being sampled for body orientation, engaged activities, voice and EDA measures using multiple cameras, microphones and wearables. Their early experimental findings revealed the potential of this dataset in terms of driving multimodal activity recognition when the interactions catered towards children with ASC. Likewise, Liu et al. offers a dataset named "Response to Name Dataset". They use this dataset for evaluating interpersonal communication abilities and lack of response skills in children with ASC using the following modalities: speech recognition, face detection/alignment, head pose estimation, the response speed, eye contact duration and head orientation [Liu et al., 2017]. Marinoiu et al. introduced a large scale multimodal dataset where they present action and emotion recognition tasks recorded during robot-assisted therapy sessions of children with ASC [Marinoiu et al., 2018]. Overall, these datasets are insightful for studies around children with ASC but they do not take into account the interpersonal social interaction between children with ASC and non-ASC. In this regard, our research (as explained in section 4.3.2 Data sharing) has defined a data resource for making all collected data available to research.

Chapter 3

RESEARCH FRAMEWORK

3.1 Introduction to this chapter

In this chapter we discuss the materials and methods being used for developing structured evaluation methods for full-body interactive technologies, where children with ASC practice SIBs. As a first step, we present the full-body interactive environments being tested in this thesis. We present Lands of Fog, a large-scale full-body interactive mixed reality environment, which was designed and developed previously by FuBIntLab, along with its newer version which we developed during this thesis. Next, in order to compare the potential of the new version of Lands of Fog with a typical intervention used by therapists, we present an adapted version of a commonly used intervention based on LEGO bricks that we designed in this thesis. We also discuss in detail the design and development process, physical interface, game mechanics and interaction design used for all of these environments. After this, we present the data acquisition stage used for this work with the type of data collected, definitions to this data, and the instruments that we developed and used. Finally, we explain how we collected this data in the experimental design stage.

3.2 The Environments

3.2.1 Lands of Fog

The Lands of Fog is an installation based on a virtual environment, specifically developed as a space to encourage social initiations in children with ASC while playing with children without ASC, where exploration and discovery of hidden virtual objects and surprises is possible. It is an ecologically valid face-to-face full-body interaction environment in which a natural dynamic of collaboration and exploration is fostered.

3.2.1.1 Motivation

Lands of Fog was developed as part of the research project “IN-AUTIS-TIC: integracio social dels nens amb autisme a travs de les TIC” (Social Integration of Children with Autism through ICT) and was funded by the grant RecerCaixa 2013. The project was developed by the Full-Body Interaction Lab (FuBIntLab) at the Universitat Pompeu Fabra in Barcelona in partnership with lead psychologist Pamela Heaton from Goldsmiths (University of London) and with psychologistst from the Special Unit for Developmental Disorders from Hospital Sant Joan de Déu.

The primary goal of the research project was to develop a Mixed Reality full-body interaction environment designed specifically for children with ASC as a space to practice social initiation with a non-ASC peer. It is aimed to design many different experiences for the same face-to-face full body interactive configuration. Each experience can address different goals for the children and therefore the therapist on the long run can have a good repertoire of experiences which can be selected as they are needed by an intervention program. Another possibility is to define different goals for every child to adapt the experiences to the specificity of every child. This builds from the interest ASC children show towards video games, hence they begin with a positive attitude towards such experiences. Furthermore, as they are adept in the field of video game logic, these children approach the systems with comfort, competence and empowerment.

The other purpose of this project was to bring ASC children together with the non-ASC children, on the one hand trying to help Autism community to better cope social integration challenges faced within society and on the other hand trying to make non-ASC children see autism in a more natural way and help them integrate in the population. Lands of Fog is an ecologically valid environment in the sense that it defines a similar situation to that of meeting other children in a school playground or a public park. The children can approach the other children in similar scenarios where they can see each other as valuable play partners that could help to develop and maintain friendships.

Another goal of this project was to define context for interaction of users; i.e. it mediates experience as a context, as opposed to defining a virtual character that confronts the children, tries to lead them, and mediates the activity. This way of defining a context rather than a guide that mediates experience, also frees the therapist from having to mediate experience. This way, the therapist can concentrate on observing the children at a distance and better understand their behaviors. In this way the therapist can help them if deemed necessary. The therapist can even become one more participant if desired and play along with the children without having to worry about the coherence and flow of the experience. Moreover, the system can also be adaptable in real time depending on the behavior of the children and their interaction. Therefore, the system can sense the real time status of the interaction and try to adapt the context of the play session in the best possible way.

3.2.1.2 Background

Initially, Lands of Fog was inspired by a full-body interactive artistic installation called "El Ball del Fanalet Lightpools" [Parés and Parés, 2001]. In this installation users could use hand-held pointers shaped as paper lanterns to explore interactive objects (props) within the projected virtual environment. These lanterns were crafted in the shape of Catalan "fanalets", a culturally symbolic artifact used in popular dances. The main goal of this installation was to explore, back in 1998, the social possibilities of Virtual Reality as a medium.

In 2001, during the brainstorming stage of the EU-funded MEDIATE project (the pioneer project of interactive multimodal full-body interaction spaces for children with ASC) the psychologists of the consortium organised an informal gathering of children with ASC to observe how they behaved while experiencing Lightpools, as this installation was being exhibited at the "National Museum for Photography, Film and Television" of the UK. This experience demonstrated the potential impact of the installation in fostering SIBs in children with ASC. This unplanned fortunate discovery paved the way for designing new experiences for children with ASC using some of the interaction design techniques employed within Lightpools.

Along with the goals defined in section 3.2.1 Lands of Fog, lessons learned from the Lightpools project were framed together with the background knowledge of interaction design, with the outcomes of the participatory design workshops, and related technical considerations for the development process of Lands of Fog.

3.2.1.3 Setup

The visual interface is configured as a circular projection on the floor, six meters in diameter, generated by images from two Full HD projectors to achieve high quality and immersive visuals (Figure 3.1). The surround sound system also contributes to the immersiveness of the user. Despite its immersive nature, the system allows the users to interact in a non-invasive and non-encumbered manner, providing them the opportunity to interact with the virtual environment in parallel with the physical environment. In particular, the circular projection facilitates movement by the absence of corners, inviting players constantly into the main playing field. In addition, the diameter of 6 meters provides enough room for children to wander freely and to gain casual exploratory behaviors within the environment. In this large-scale MR environment, the experience of personal space is not invaded and users are allowed to interact with each other at their own pace.

The children interact with the projected VE through a physical object acting as a pointer (with the shape of a butterfly net) (Figure 3.2). This object allows them to better focus on the activity (acting also as a cognitive offloader) and allows them to have a better sense of control over the VE. It has been also mentioned that the visual and tactile properties of tangibles may help focus children's attention on social interaction and allow children to create a common context in terms of sharing ideas and objects [Marwecki et al., 2013, Villafuerte et al., 2012, Farr et al., 2010, Ortega et al., 2015]. This notion was also taken into consideration in the implementation of the pointers. A multi-camera system tracks these pointers on the playing field based on a ring of LEDs that provide a cicle of colored light that the system identifies (red or blue).

The tracking system of the pointers that the users hold was custom developed by FuBIntLab using OpenFrameworks, OpenCV, and TUIO in C++. The MR application was developed with Unity game engine using MiddleVR to manage the double projector system that forms a single HD image (1920x1920 pixels).

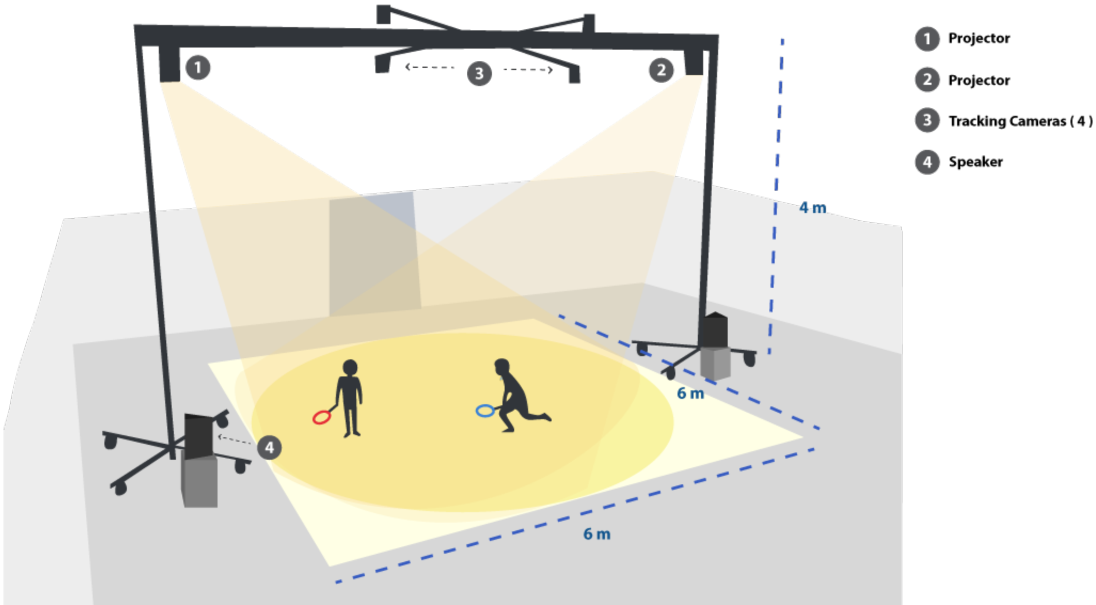


Figure 3.1: Diagram of the Lands of Fog system

3.2.1.4 Content

The game is based on an imaginary world designed through Participatory Design Workshops together with four male children with ASC (10-12 years old). Children proposed to have different zones in the virtual world. In this regard, four different environments (as biomes of ice, lava, forest, and ruins) were designed that converge into one virtual environment (Figure 3.3). Children also defined that each zone would be populated by unique objects and creatures such as a Yeti, a Golem, a Coral Girl, and a Crab Man (Figure 3.4). As a result, a total of 14 creatures were designed that could be discovered, and each child could only have one creature at a time. The creatures were also associated with the different regions of the environment and in each region there were different objects to be discovered. Moreover, a total of 16 different objects were designed and these objects proposed to be interactive and could be transformed with different sounds and animations [Mora-Guiard et al., 2016]. The game scenario also included hunting butterflies in this virtual world which led to design of the pointers in the shape of butterfly nets.

3.2.1.5 Interaction Design & Game Dynamics

Following the feedback from the participatory design sessions, the range of sounds and visuals, the changes in the characters and objects, as well as the surprising elements, all generated by the system, designed to provide a positive reinforcement to players and encourage creative exploration, which would have otherwise been provided by the therapist, having to mediate, guide and support the progress of the children. Moreover, the Lands of Fog condition is designed to provide structure and assistance to therapists to observe and mediate the session externally, while ensuring a stable and predictable context for children to put social interactions into practice.

The Lands of Fog condition designed provides a context for the two players to adopt ex-

Figure 3.2: The children interact with the projected VE through a physical object acting as a pointer (with the shape of a butterfly net which is illuminating)

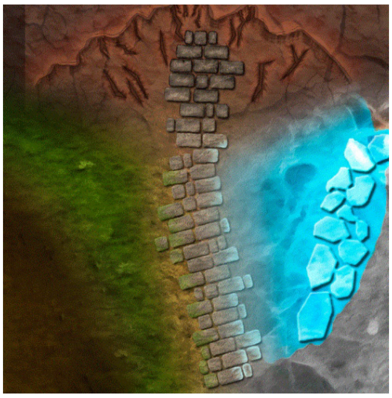


Figure 3.3: Four biomes were present in the environment: ice, lava, forest, and ruins.

Figure 3.4: Children could merge creatures by collaborating with a partner, to explore various possibilities.

ploration and discovery attitudes by navigating physical and virtual space simultaneously in a face-to-face configuration. Discovery of interesting situations is meant to engage the children and generate the sufficient internal drive in the ASC children to want to communicate with their peers. The full-body interaction permits them to use not only verbal communication but also facial expressions and body language. The configuration also allows for joint attention and collaborative activity. A detailed explanation of the interaction design elements in the game will be explained in the next sections based on two versions of the system developed which contributed to the final format which was evaluated in this thesis.

3.2.1.5.1 Lands of Fog v1

Design of a Virtual Fog: In order to motivate user engagement and exploration it was decided to allow the users to see only a small portion of the large environment at every one moment. This was achieved by covering the virtual world with a layer of virtual fog which opens only a small hole in it at the position where the users place their nets. This interaction design strategy was taken from the Dalsgaard and Dindler's [Dalsgaard and Dindler, 2014] study which they describe as creating a "peephole" in the scenario and has shown to be a good practice for promoting exploration and user engagement.

Hunting Fireflies: When placed in unknown environments, children with ASC might adopt passive attitudes. In order to tackle this and provide a simple introductory game mechanic, as a first "low floor" [Papert, 1980] contact, the system is designed to present swarms of fireflies which move through the virtual fog and users are allowed to hunt these fireflies. These swarms move in random trajectories, emitting a bright yellow light, and attracting the user's attention, above and below the fog. The aim is to get children curious about these swarms that are appearing and disappearing and motivate them to catch a set of two fireflies with their colored nets. On catching two fireflies, these change color from yellow to the color of the childs net (red or blue) and they start to follow the child wherever s/he goes. When, for a short period, the child shows no interest in these swarms, the system subtly shifts their trajectories, gradually bringing them closer to the child, making them to be easily recognized and accessible, even up to a point in which the child can hunt some of the fireflies in the swarm almost without effort. This subtle system reaction can be seen as an additional assistance given to the child without stigmatizing him/her. In other cases, users have to bring their nets close to the swarms to hunt the fireflies. Whenever a user catches one batch of two fireflies, it was decided to make the swarm move away from the user to keep fostering an exploration attitude in the children. After hunting a certain amount of fireflies, the set of fireflies that already belong to the child transform into a creature (from a repertoir of 14 different ones) that becomes a virtual partner for the user. If users continue to hunt fireflies, after a certain new number have been caught, the appearance of their creature changes (it is the same character but with different colors or textures). All these system actions were aligned with the goal of maintaining user engagement with the system and it was following the design principle presented by Davis et al. and Jordan [Davis et al., 2010] which is about beginning interaction with simple objectives and transitioning to more complex objectives

Reducing Competitive behaviors: Rather than leveraging competitiveness as a promise of engagement and motivation in gameplay, it was decided to embrace children to adopt a more positive mindset with plot strategies aimed at creating a sense of empathy between the children and their creatures, and teaching the children about the positive effects of collaboration. There was a possibility that when the children were hunting fireflies, they would attempt to steal their partner's fireflies or fight for the same firefly swarm. To prevent this, different design strategies were employed. The first strategy involved engineering the fireflies to remain uniformly distributed around the virtual environment, ensuring that there was no moment where the swarms clustered together. This way, children would not be tempted to interact with the same group of fireflies. The second strategy was to change the body color of the fireflies to the color of the pointer of the player after they had been captured. This made children realize that those fireflies now belonged to their personal swarm and were unlikely to be hunted by the other peer. The other strategy was to portray the swarm of gathered fireflies as a rising uncountable cluster not allowing children to compare their collected number of fireflies. Moreover, the game was originally designed to adopt an open-ended format of play in which game rules or goals are not defined allowing users to continue playing in a repeated exploratory manner without competing with each other to achieve predefined goals (Figure 3.5).

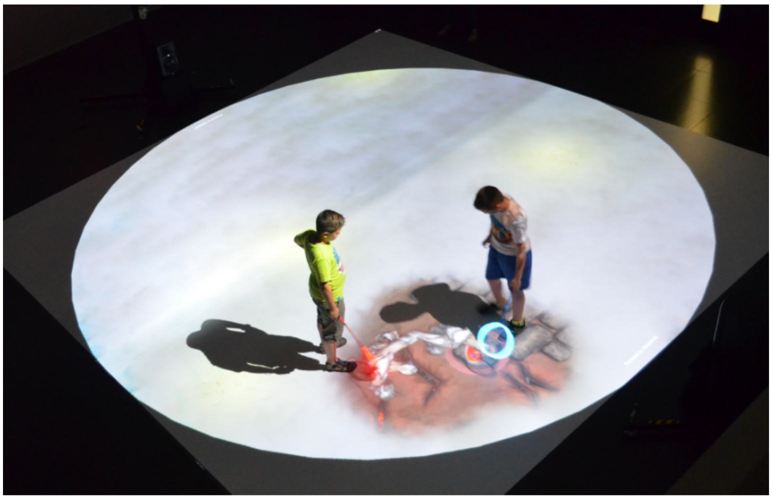


Figure 3.5: The lit circles signify swarms of fireflies moving just below the fog. The user sees the world below through the hole in the dense fog.

Creature Feedback: When adopting the strategy of having virtual partners, it was planned that creatures could be suitable for subtly bringing users together and softly mediating shared opportunities. It was decided that when the users explore through the virtual world, creatures must accompany the users, so that the children could form a sense of empathy and companionship with their creatures. Moreover, the characters were able to communicate with their users directly. For example, after the user stays in the playing field with a passive attitude, it was decided that the character of the participant would look up towards the user to catch attention and create engagement. Moreover, characters were able to point at virtual objects when the user is nearby that object to help the users discover the interactive nature of that object. Additionally, when users' creatures come close to each other, they perform a greeting action towards the other. Therefore, creatures are models of social behaviors for the players, and also try to help children understand that there might be potential interactions between virtual elements.

Encouraged Collaboration: Many studies have shown that play in children with ASC is often solitary [Bauminger et al., 2010, Kasari et al., 2012, Symes and Humphrey, 2011]. In this regards, when designing Lands of Fog, it was planned to leave a room for children to do solo activities so they could have their own unique game experience while they still do not feel comfortable to approach the other users. When they figured out the fundamental dynamics of the game, it was decided to slowly incorporate the notion of collaborative playing with their peers. Nonetheless, we wanted to adopt what we call an "encouraged collaboration" approach [Crowell et al., 2019a] rather than the "enforced collaboration" approach [Ben-Sasson et al., 2013], which means that children were not forced to collaborate in the game. Instead, the game offered fun incentives such that the children would choose to work on their own will. For example, when only one participant had a creature, that participant was able to duplicate and share his creature with the other user by moving his net close to his peer's net, constituting an easy interaction and thus simplifying the understanding that things can be shared. Moreover, when both users had a creature, and when users put their creatures together. users would discover new game mechanics that encourage them to play the game collaboratively. For example, when users brought their creatures together, these would merge and morph into two new creatures through a ritual where they jump towards one another. Therefore, if the children decided to explore all the creatures that populate the virtual world, they had to have a joint plan and work together to merge their creatures. Another mechanic was to make the virtual objects found in the virtual world interactive in a way there was an animated response from the element when both children brought their creatures together nearby this virtual object. By using this audiovisual feedback as a game reaction, the children could understand that when during collaborative play they can expect more thrilling and surprising gameplay than solo play.

3.2.1.5.2 Lands of Fog v2

Following the proposal from the team in terms of developing a second version of Lands of Fog and comparing its potential to that of a typical intervention used by therapists, few changes were added to LoF v.1. Since the timeline of this task was aligned with the start time of this thesis work, this thesis work also contributed to the development of these changes. These changes are mainly focusing on fixing the flaws found in the LoF v.1 in terms of slightly improving the setup, content and few interaction design elements. First of all, the design of the handheld pointers were improved to be more robust, as we had found that the initial nets were easily broken due to rough movements. Moreover, the design of the new nets included a more elegant approach keeping the batteries and cables more hidden. Additionally, it was decided to use LED lights to cover the net instead of electro-luminescent wire which offered increased visibility to the tracking cameras. Secondly, we had seen that during a 15-minute play session, the children could discover much of the virtual world. In this regard, in order to improve the open ended nature of the interaction, we added a second map with additional interactive elements. Since the first map was designed by a group of male children with ASC through participatory design sessions, a female took part in the design of this second map, which was a beach scene. We designed a mechanism through which the children could swap from one map to the other and back again. The transition between maps was triggered by both children manipulating an interactive object (prop) in the shape of a bottle with a map inside. Moreover, we added a door in the center of the virtual world which could be unlocked by bringing together two halves of a gold key which were randomly hidden inside the virtual elements and could be only collected through manipulating these props. Finally, bug fixing and fine tuning processes took place with the speed of the char merging and prop manipulating actions, and with the appearance of the fog and some of the characters.

3.2.2 A traditional play therapy: LEGO

A traditional play activity was needed to compare the effect of Lands of Fog condition on the behaviors of children with a typically used strategy in real world interventions.

The strategy we chose to implement was created based on Daniel Legoff's well-known therapy. Legoff obtained promising results in improving the acquisition of social skills in children with ASC especially in peer interaction scenarios [LeGoff, 2004]. We have adopted this strategy and created a customized LEGO condition, with the advice from the team of experts that we collaborated with in the Hospital Sant Joan de Déu. It matches the game dynamics defined in the MR environment such as the ambulatory nature of the activity, the exertion which children have to go through, standing position with downward looking attitude, scenarios around proximity between children. This is achieved by a specially designed LEGO play environment in the form of a hexagonal-shaped table that allows free movement around it, as opposed to playing by sitting on the floor. The children start the game by looking for pirate-themed figures in buckets that contain the LEGO blocks, placed on the six vertices of the table. Following the discovery of these figures they continue by collaboratively building a ship for the figures in the center of the table. A notable difference with the Lands of Fog condition is that this one is a passive system in which the therapist mediates the activities (Figure 3.6).

Figure 3.6: Children work together in building constructions in LEGO condition

3.3 Data Acquisition

3.3.1 Wearable Multichannel Psychophysiology

[Parts based on the article Sayis, B. Pares, N. Superhero-collar wearable for children with ASC for acquiring reliable physiological signals. (under preparation)]

3.3.1.1 Motivation

The use of wearable sensors is a promising research direction for monitoring psychophysiological states in individuals with ASC [Picard, 2009]. However, when it is decided to use a wearable technology in ASC research particular attention should be paid to attributes such as comfort and unobtrusiveness [Koo et al., 2020, Taj-Eldin et al., 2018]. Problems arising from these attributes can contribute to a child's unwillingness to wear additional devices. Moreover, physiological signals are prone to contamination by motion artifacts [Boucsein, 2012]. Particular consideration should be paid, especially in full-body interaction environments, to using both ECG and EDA, since they are susceptible to motion artifacts due to movement of the peripheral body. In this regard, we designed and developed a wearable specifically for children with ASC, optimizing the monitoring of pschophysiological signals in full-body interaction environments. This wearable was designed to make the experience fun and acceptable by the children. It does not restrict movement of children as hands are free from sensors or cabling, which allows them to interact more naturally within the full-body interaction environments.

3.3.1.2 Background

The desire to track individuals reliably and over long periods of time has generated interest in comfortable wearable technologies, unobtrusive instruments that can be worn to capture psychophysiological data during normal daily life. Many of these projects have been driven by MIT researchers, creating technologies that can be worn comfortably during daily activities [Picard, 2009]. Some factors that have been emphasized in the development process of these wearables also included easy to put on or take off nature of the wearables. Their collaboration with people with ASC to co-design toolkits that customize the control and output of their devices resulted in a wireless EDA, HRV (PPG based), temperature and motion sensor packaged on the wrist. However, their choice of sensor and electrode positions and types could be argued as they play a very important role in the quality of signals.

EDA is typically measured from the palms of the hands and fingertips[Neumann and Blanton, 1970]. However, in a recent study by [van Dooren et al., 2012], the responsiveness of 16 different recording positions of EDA was quantified and it has been shown that foot, fingers and shoulders are most responsive, whereas arm, back, armpit, and thighbone are least responsive. On the other hand, when it comes to measuring HRV, the most commonly used and effective method that researchers have used for decades is ECG. With electrodes, ECG can be acquired at almost any location on the body. Best results, though, can be obtained with proper placement of electrodes on locations far from high muscle activity such as around the chest on a rib or under the clavicle. This configuration is referred to as LEAD II, based on Einthoven's terminology [Wasilewski and Poloński, 2012]. On the other hand, Photoplethysmography (PPG) is another sensor technology to measure HRV. PPG sensors use a light-based technology to detect the blood flow rate as driven by the pumping action of the heart. However, while PPG has been a common option for many wearable devices, perhaps due to its lower production cost and ease of use, it does not always provide accurate data. PPG recording challenges include interference of ambient light in the recording site and variety of skin color which has a certain effect on the reflection rate of the PPG light sensors. Moreover, PPG operates better on body parts that have a high blood vessel concentration, so it can be difficult to receive accurate signals from the wrist [Bent et al., 2020]. Additionally, due to the need to quantify the amount of ambient light and determine the compensation required for canceling its effect, PPG sensors require a relatively long settling period compared to ECG sensors. ECG sensors do not need lengthy settling periods, so meaningful readings can be obtained, rather shortly after initialization or any intervening movement artifacts. For example, the sensors used in MIT research are placed on the wrist and movement of the arm has the potential to create a strain on the sensor connection. Besides, they are using dry electrodes for their EDA sensors although the use of cup electrodes mounted with adhesive pads and conductive gels is one of the best ways to help reduce the impact of electrode movement on signal integrity [Fu et al., 2020].

The patent application US 2013225966A1, whose applicant is Macia Barber et al. [Barber, 2013], presents a robust wearable design for acquiring physiological signals in high physical activity scenarios. This wearable is based on a shirt-like garment which includes sensors hidden in an improved silicone rubber placed on the chest area of the garment. However, such a wearable might present a major drawback in terms of acceptability by children with ASC. Although the specified sensing device seems to be non-invasive, it might not be easily put on or taken off by children with ASC and this process can lead to extra stress or anxiety. Moreover, approaches for physical comfort and anxiety reduction in ASC also suggest that using a weighted blanket around the shoulder might help reduce anxiety [Denny et al., 2018].

Thus, from this knowledge in the state of the art, it is clear that there is still great interest in the development of a wearable comprising: adequate electrode and sensor types; adequate placements for reliable acquisition (especially in movement); as well as comprising garments providing aesthetics and physical comfort, especially for children with ASC.

3.3.1.3 Design & Development

Our design allows different types of biosignal acquisition systems to be used within the design. In our design we are specifically using the Biosignalplux system [PLUX Wireless Biosignals S.A.,]. Biosignalplux (figure 3.7) is a ISO 13485 class medical device for biosignals' acquisition and wireless transmission in real-time up to 3000Hz with internal memory acquisition option. Biosignalplux integrates a microcontroller, a module for power conditioning and battery management, internal memory, a wireless communication module that uses Bluetooth technology, allowing it to be connected to a computer, mobile phone or any other device that includes a Bluetooth receiver. The research kit can work with a set of physiological sensors: ECG, EDR, ACC, EMG, EEG, BVP, ST, and PZT. In our design, we are specifically using the following biosignalplux sensors: Electrodermal Activity (EDA), electrocardiogram (ECG), and Accelerometer (ACC). We also considered various choices of types of sensor and electrode, and their location on the child's body since they play a very important role in the quality of psychophysiological signals. We used gelled self-adhesive disposable Ag/AgCl electrodes (Figure 3.7), as adhesive pads are another way to help reduce the impact of electrode movement on signal integrity.

Previous findings on EDA measure indicate feet, fingers and shoulders as being the most responsive. We therefore chose to place EDA electrodes on the shoulder (Figure 3.8) in order to get a reliable signal which is also least affected by movement artifacts. On the other hand, reliable ECG signal results can be obtained when both measuring electrodes are placed on locations relatively devoid of underlying muscle. Therefore, to avoid movement artifacts, electrodes may be placed on the torso. We placed two of the three lead ECG electrodes on the upper region of the chest and placed the ground electrode on the

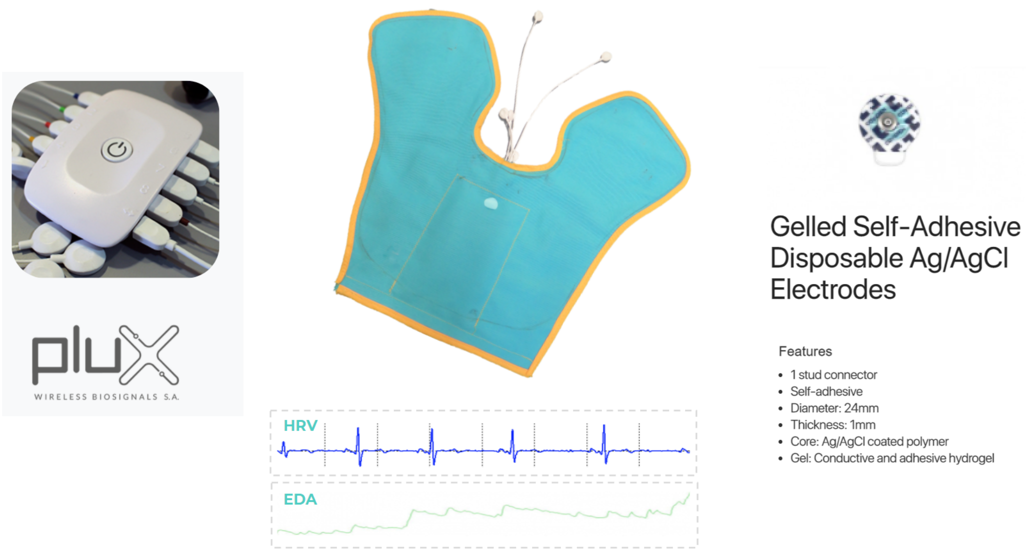


Figure 3.7: Views of biosignalplux research kit (left) / Superhero-collar wearable (middle) / gelled self-adhesive disposable Ag/AgCl electrode (right)

spine section of the neck (Figure 3.8). Furthermore, the position of the ACC in the wearable was aligned with the shoulder EDA electrodes so that the movement artifacts on the EDA signal could be related to ACC data.

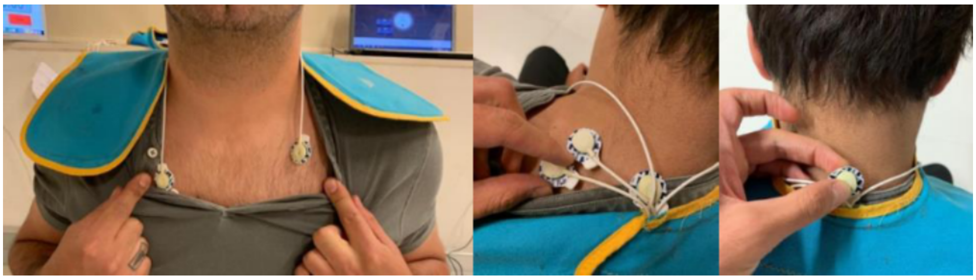


Figure 3.8: Position of EDA and ECG electrodes.

The wearable form factor design that we chose was that of a "super hero's cape" to make it, not only acceptable for the children, but also motivating and fun. It holds the sensors and cables at the same level and hides them as much as possible, while providing necessary ergonomic requirements. Under the blue cover seen in (Figure 3.7), we are using a Hair Cutting Collar Shoulder Cape as seen in (Figure 3.9). The collar's front weights (metal circles within the front tips) help balance the weight of the sensors and the main sensing unit, which are located at the back. On the other hand, approaches for physical comfort and anxiety reduction in ASC also suggest that using a weighted blanket around the shoulders of ASC children (Figure 3.9) can help reduce anxiety [Erickson et al., 2012, Denny et al., 2018]. We therefore designed our wearable so that the weight that our wearable puts on the shoulders can also be seen as an anxiety reduction factor. The choice of the colour, blue as the main color and yellow on the edges, was also studied to give the relaxing and superhero costume-like feeling.

Figure 3.9: Hair cutting collar (left) [CoolBlades, 2021] / A weighted blanket for autism (right) [Sensetoys, 2021]

3.3.1.4 Verification

As a first step, in order to understand whether the design and construction of the wearable was headed in the correct direction, we conducted verification studies around this configuration. We first tested the reactiveness and robustness of the suggested electrode positions in terms of ANS activity and movement artifacts.

We chose to place EDA electrodes on the shoulder in order to get a reliable signal which is least affected by movement artifacts (especially compared to hand movements during play) [Kasos et al., 2020, Posada-Quintero and Chon, 2020]. However, as this approach has not been common, we wanted to compare this location's effectiveness with a configuration where EDA electrodes are attached to both palm and shoulders. In this regard, we simultaneously recorded EDA signals from both locations and asked our lab members to go through a short procedure to test this configuration (Figure 3.10).

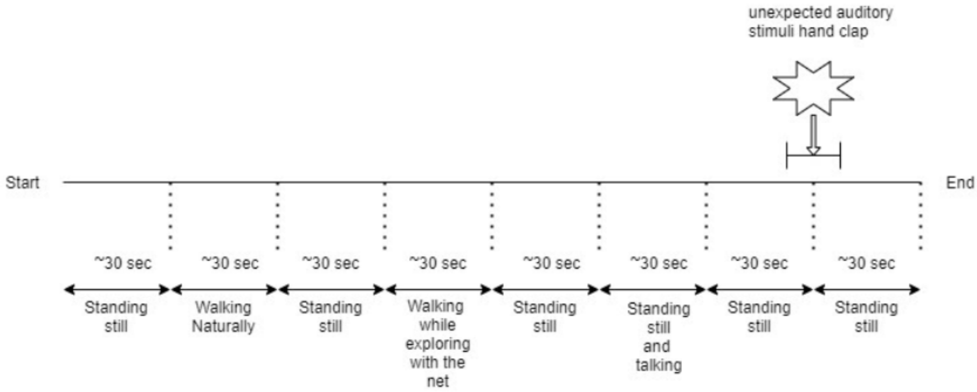


Figure 3.10: A short procedure to test the initial configuration of our wearable prototype

We adapted the procedure that Macia Barber et al. [Barber, 2013] used while they were testing the robustness of their wearable for daily activity scenarios. We also considered the possible activity range that could take in our experimental conditions such as exploration using the butterfly nets in the playing field, walking naturally and talking. We also added an unexpected auditory stimuli (hand clap) in order to test whether the shoulder positioning is as effective as the palm positioning in terms of detecting a possible ANS related activation. When we applied visual inspection on the raw signals, it was possible to see a reaction to the unexpected auditory stimuli from both EDA electrode positions (Figure 3.11).

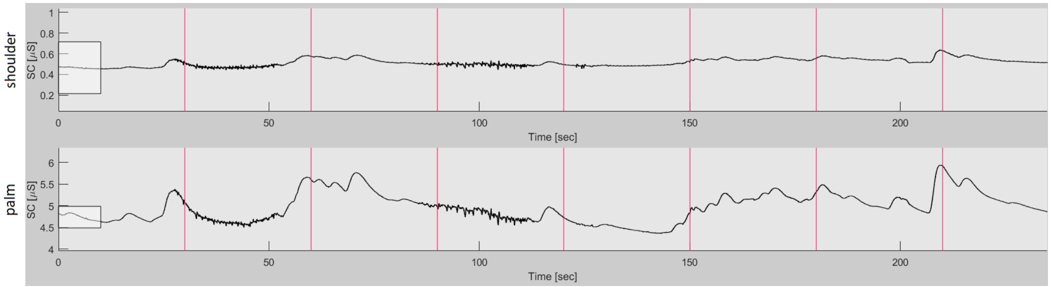


Figure 3.11: Reactions to the unexpected auditory stimuli from both eda electrode positioning, shoulder(top), palm(bottom)

The amplitude of peaks in the skin conductance responses and the initial skin conductance level from the shoulder were indeed smaller than the values measured from the palm. This might explain why the palm is chosen as a commonly used approach. However, it should be noted that the movement artifacts associated with the two positions during walking and exploration were relatively smaller in the shoulder compared to the palm.

These preliminary findings were promising and opened the door for us to validate the current configuration of our wearable in our pilot trials with children with ASC.

3.3.1.5 Validation

In order to understand whether the wearable is adequate for children with ASC and at the same time provide robust and meaningful measurements, we started running a validation study within the context of our experimental design (section 3.4 Experimental Design). This study was undertaken with 8 children (n=1 female, n=7 male), aged 8-12, from an inner city school in Barcelona as a part of a school trip to FuBIntLab. Four of these children were high functioning ASC. We ran four trials involving pairs of children with ASC and non-ASC. In each trial participants experienced Lands of Fog for approximately 15 minutes and wearing their respective wearables. At the start of the trials, the children looked at this wearable with some curiosity, and after researchers showed themselves how it could be worn without harm, children did not mind wearing it. Most children did not express discomfort at wearing the wearable, but some were initially nervous about removing the adhesive pads. After the trials, we visually inspected the physiological signals collected in terms of movement artifacts and detection of possible ANS related activity. We could see that the wearable did allow for collection of reliable data. One example worth mentioning from one of these trials was one of the ASC participants' reaction to a specific game event taking place during Lands of Fog. When the game dynamic hunting insects mentioned in section 3.2.1.5.1 Lands of Fog v1 led to the climax with the caught fireflies magically fusing into a character, this participant gave a startle response by jumping back from the event (Figure 3.12, Figure 3.13). We could see this reaction clearly from our physiological measurements in terms of ANS activity and movement artifacts (Figure 3.14).

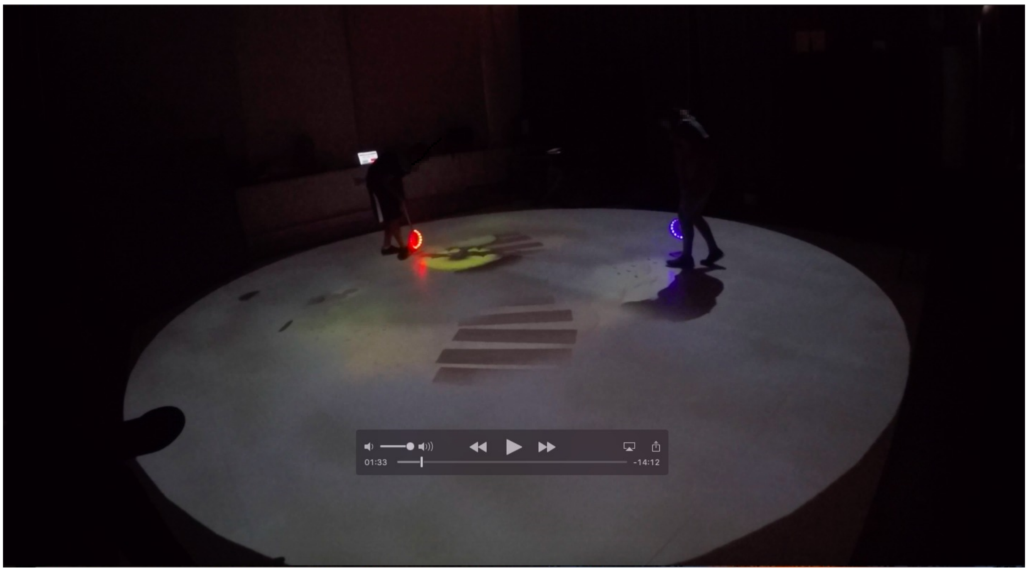


Figure 3.12: The moment child with ASC (holding a red net) sees that the character is created in the VE (with yellow glowing light around it), 1 s after the startle response begins by jumping back.

Figure 3.13: The moment participant's startle response represented by jumping back action take place, startle response ends.

Typically for an event related SCR, the latency period between stimulus onset and the first significant deviation in the signal is between 1-5 s [Posada-Quintero and Chon, 2020]. In our case, 1 s after the ASC participant (holding a red net) saw that the character was created on the floor, the SCR response with an approximate amplitude of 0.7 microsiemens could be seen in the EDA signal (Figure 3.14).

Following this moment, the participant's startle response represented by jumping back action started taking place. This resulted in deviation in ACC signals (X,Y,Z directions) as well as in ECG signal. Although this jump was relatively strong, we could see that both EDA and ECG signals were robust enough to be detected clearly. Although these pre-

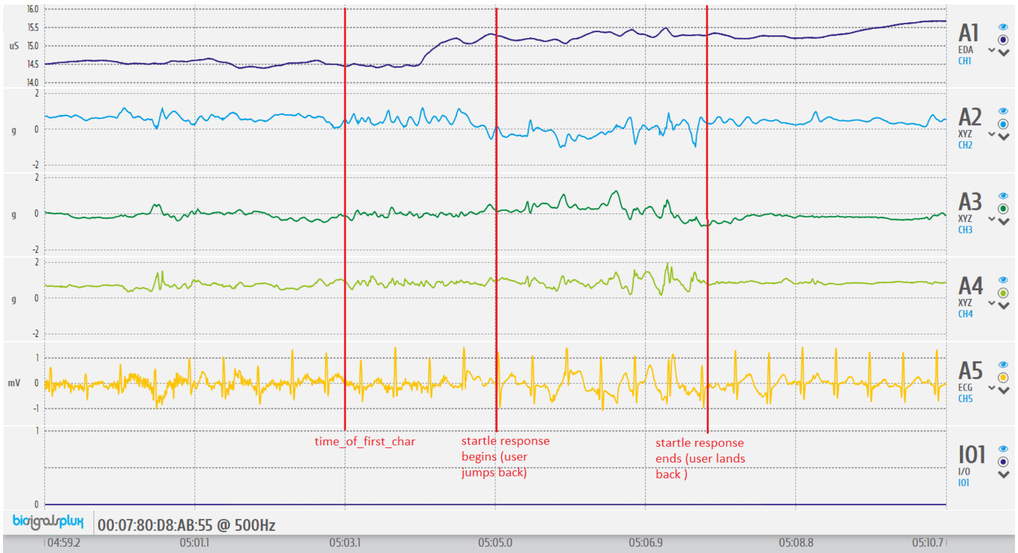


Figure 3.14: Opensignals [Osr, 2016], an open-source software to check the signal quality, view of raw psychophysiological signals in a 10 second window. Events are marked with red horizontal lines (time_of_first_char, startle response begins, startle response ends)

liminary results gave us confidence in terms of readiness of our wearable for starting the experimental trials, in the EDA signal of one of the participants we detected an extremely high initial level of SCL. Our investigation on this matter led us to link the problem to the setting of the laboratory temperature and pre-state of the participants. These trials had taken place during the summer months in Spain, and the children arrived at our lab having perspired from walking outside. Moreover, it should be noted that thermoregulation may play a bigger role in EDA measurement location, when comparing shoulder electrode placement to other parts of the commonly used EDA measurement locations (i.e. finger palm). In this regard, we decided to control the laboratory temperature more carefully in our experimental trials.

3.3.1.6 Acquisition

In our experimental trials, the psychophysiological signals were acquired in an "offline mode" to the internal memory of the biosignalplux devices at the sampling rate of 500 Hz and digitized with a resolution of 16 bits per sample. It should be noted that "offline mode" wording here is based on the terminology used in the biosignalplux device's manual [Osr, 2016]. It should not be misunderstood that the psychophysiological measures taken in this work is based on the measures taken after the play sessions. The psychophysiological measures were continuously taken during the duration of the experimental session and they were being recorded locally, in real time to the internal memory of the devices. When the whole experimental session was over, we transfered the data from the biosignalplux devices to our computer. The total duration of the experimental session was approximately 1 hour and 15 minutes (see section 3.4.3 Procedure). On the other hand, although biosignalplux devices allow bluetooth acquisition, this communication option was proven unreliable with too many connection losses when recording the signals simultaneously from two biosignalplux devices over an hour period. As a result, we chose to record the psychophysiological signals in real time to their internal memory and transfer the data to our computer after the experimental trial . On the other hand, we used bluetooth communication through the open-source software OpenSignals [Osr, 2016] to configure the device, to check the signal quality and correct attachment of the electrodes before starting the experimental trials, and to finally download the recorded data stored in the internal memory.

3.3.2 Video Recordings

All sessions were video recorded via two GoPro HERO4 Silver cameras that provide a wide angle view of the whole interaction space, stationed at opposite sides of the play arena (Figure 3.15). We used external high quality microphones (Rode VideoMic Go) connected to the GoPros to acquire good quality sound in the video recordings since the conversations are crucial to understand and code the SIBs. We recorded the sessions with 30 fps in Full HD resolution (1920 x 1080).

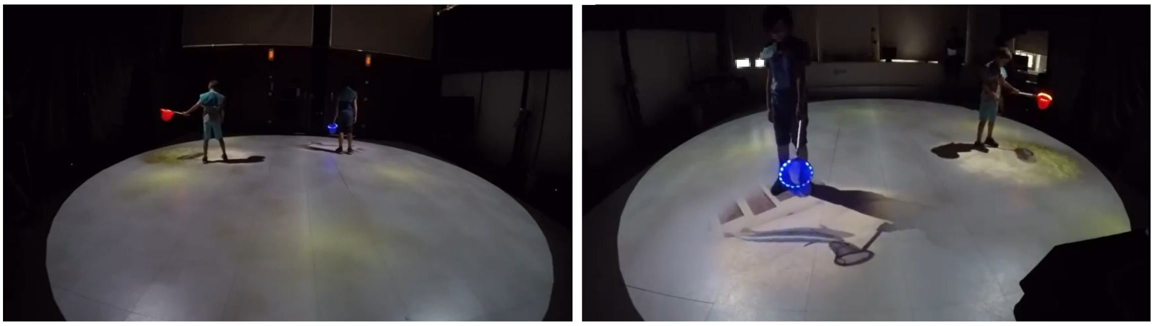


Figure 3.15: GoPro cameras recording points of view from Lands of Fog play session to be used in video coding: position 1 (left) and position 2 (right)

3.3.3 System Logs

In the development of Lands of Fog mentioned in section 3.2.1 Lands of Fog, one principal goal of the system was to detect behavioral features of the participants during interaction and have the system respond accordingly. As the environment employs sophisticated techniques, inviting and motivating the children to interact with the game in terms of modifying objects and merging characters in a collaborative way, these specific events have been recorded to the system's log file for further analysis after sessions. The log files are in a time series data format, including position data for every collected frame. The log files also include a timestamp for each specific event taking place during the session. Through these log files we were able to have a general overview of each session. The definition of each specific event being logged can be found in Table 3.1 (aligned with the definitions found in section 3.2.1.5 Interaction Design & Game Dynamics).

Table 3.1: Definition of the each system log

System Log	Definition
hunted_insect	When users bring their nets close to the flocks of fireflies, two insects are caught. This is represented by changing the fireflies' color to match the user's butterfly net color while the rest of the flock moves away from the user
character_created	After hunting a certain amount of insects, it was decided that fireflies would transform into a creature that would become a virtual partner for the user
character_changed	In order to maintain user engagement with the system, as users continue hunting insects their creature changes appearance. This mechanic was implemented to foster the children's sense of exploration as they discover different versions of the creatures
character_points_at	Having the character pointing at something was designed to help the users discover the nearby elements which were interactive, hence it is activated when the character of a user approaches the virtual elements. At that moment, the user's creature points towards the element and makes an exclamatory remark
character_being_idle	After staying in the playing field without any action, it was decided that the character of the participant would look up towards the participant to catch attention and create engagement
manipulate_prop	The virtual elements can only be activated when both users bring their characters close to the element at the same time. When both creatures converge by the object, they will manipulate the virtual element, which will respond with a unique playful animation and then disappear, while the creatures celebrate their discovery
character_merge	When users brought their creatures together, these would combine and create two new creatures, which replaced the old ones. The merge were represented by both creatures jumping towards one another and transforming into a novel creature. Therefore if the children wanted to discover all the creatures that inhabit the virtual world, they had to plan and work together to merge their creatures
character_greeting	When users' creatures come close to each other, they perform a greeting action towards the other. Therefore, creatures are models of social behaviors for the players, and also try to help children understand that there might be potential interactions between virtual elements

3.3.4 Questionnaires

Given the success of sensory objects in interviewing children with ASC in previous studies [Kirby et al., 2015] and the potential of multi-touch apps [Hourcade et al., 2013], we wanted to interview our participants through an app multi-touch tablet interface. This was also seen as a means of mitigating the ambiguity of open-ended questions, which could be restricted by limited communication abilities of the children. We tested different types of interfaces, such as multi-touch sliding bars on Likert scale questions, Smiley face cards, and computer based systems, and found that children had an instant and natural grasp of sliding scales delivered through tablet interfaces. We have seen that the ease of use of multi-touch tablets compared to the mouse and keyboard interface was higher in children with ASC while using sliding scales. The tablet interface was well suited to the motor challenges of certain children with ASC and it seemed like it made answering questions less monotonous.

We adapted the Affective Slider scales [Betella and Verschure, 2016] to our tablet interface to measure arousal (arousal_level), and valence (valence_level) of each participant. Together with these sliders, we asked questions such as "How well do you know your partner right now?" (social_status) and "Would you like to get to know them more" (desire_to_know_more) (Figure 3.16). Henceforth, we will refer to this questionnaire as "questionnaire_AS", which includes the following variables: arousal_level, valence_level, social_status, desire_to_know_more.

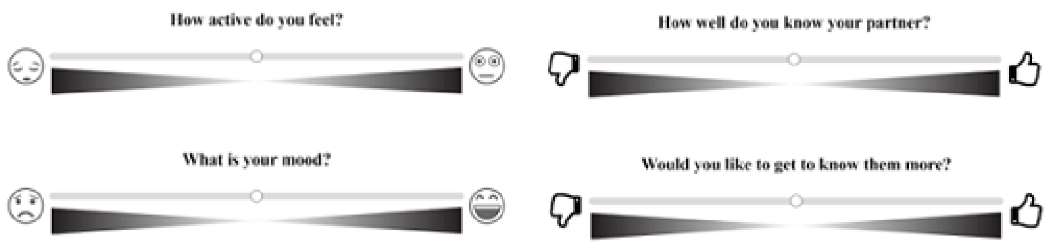


Figure 3.16: The tablet questionnaire included sliding bars and icons for arousal level, valence level (left); level of knowing the other peer, and level of motivation for knowing other peer (right).

To gather data on the child's anxiety level, we used a standardized questionnaire called STAIC (State-Trait Anxiety Inventory for Children) [Spielberger, 2010] which was also previously used with children with ASC [Lanni et al., 2012]. The STAIC is a self-report questionnaire with 20 items developed to assess children's anxiety. There are two forms, one to provide a measure of acute anxiety (state anxiety) and one to provide a measure of anxiety as a stable measure, describing the tendency to present state anxiety in daily life. Both measurements were used in this thesis work. The state part of the questionnaire was also adapted to the tablet interface so that children could mark their responses in the same interview configuration. Children answered inventory items such as "Right now, I feel calm" with one of the following responses: Not at all, Somewhat, A lot. All questions were translated into the children's first language, as we acquired them from the official version of the STAIC. Henceforth, we will refer to the state part of the STAIC questionnaire as "questionnaire_STAIC/S" and the trait part as "questionnaire_STAIC/T".

We also used the Child-behavior Checklist (CBCL), which is a commonly used screening tool for aggressive behavior, hyperactivity, and social problems. CBCL and questionnaire_STAIC/T were used to get information on each child's behavior outside of the sessions which could match to patterns we saw within the sessions. Henceforth, CBCL will be referred to as "questionnaire_CBCL".

In addition to the affective sliders, STAIC and CBCL, we also used non-standardized questions to get information about children's play experience. We paid particular attention to the use of sensory information in writing these questions such as: "Do you remember a moment when you talked to or discovered something with your partner?", which could be later related to other data sources as a strong indicator of the game experience.For example, in Lands of Fog, the moment when they merge their characters and discover the transformed new characters could be seen as a conversation starter and memorable event between participants. Moreover, the Fun Toolkit [Read et al., 2002] was also used as a model for some of the questions. For example, children were asked to report on their preference between playing Lands of Fog and a "regular" playground with friends. Henceforth, this questionnaire will be referred to as "questionnaire_Game".

All the questionnaires administered through the tablet interface were saved with the timestamps directly to a Google Drive Data Sheet (Appendix A.5). This tablet interface was presented through a custom software we developed.

3.3.5 Relaxation Exercises

Activities of relaxation such as drawing, reading and especially listening to calming music might help to calm the person. However, typical relaxation strategies are not always effective for children with ASC [Autismspeaks.org,]. Children with ASC may relax by using gentle rocking actions or engaging in a repetitive action which can include manipulating an object such as a stress ball or a Rubik's cube [Attwood, 2014].

When children with ASC are told they will be going to an unknown environment like our laboratory setting, they can get worried about this visit that breaks their daily routine. Moreover, seeing that they will wear our wearable device may also create an extra worry. In this regard, we adopted short relaxation exercises to be used in our experimental procedure following the guideline designed for relaxing children with ASC in anxiety inducing situations [Autismspeaks.org,]. The relaxation exercises that we used were based on the strategies around deep breathing, muscle tensing/relaxing, visualization and distraction activities. For further details on the technique used for these relaxation excercises see the appendix A.1.

3.3.6 Data Synchronization

Maintaining synchronization between independently recorded data sources is seen as a challenge in the use of multimodal data acquisition approaches for behavior recognition. A common approach is using a clearly defined set of synchronization steps (generated by a user or experimenter) which can be identifiable in all the data sources [Bannach et al., 2009]. This approach is also called event-based synchronization. If each data source contains a local clock, once the first synchronization action is mapped to other data streams with the help of additional events, offsets between the local clocks can be calculated and data streams can be aligned accordingly.

We followed the event-based synchronization approach and synchronized every single data source based on a universal clock presented on a screen which is placed in the experimental setting visible to all researchers. Trigger events to start and stop recordings in each data channel noted to an experimenter diary aligned with the universal clock (Appendix A.3). First data source to be synchronized was the recordings from our wearables. As our experimental setting was a multiuser setting and each participant had to wear our wearable, we needed to implement synchronized multi-device data acquisition. In this regard, we connected biosignalsplux devices with each other using the biosignalsplux sync-cables and the multi-sync splitter cable connected to a trigger push button to start offline synchronized acquisition. The universal clock and the acquisition of physiological signals started at the same time (Figure 3.17). However, in our pilot trials we encountered a technical problem with this synchronization setting that some of the recordings in the biosignalplux devices had already started without waiting for the trigger push button event. To overcome this possible synchronization problem, after starting the acquisition in these devices, we decided to push the trigger button two more times so that these events could be marked in data logs simultaneously. These trigger button events were timestamped on the experimenter diary accordingly.

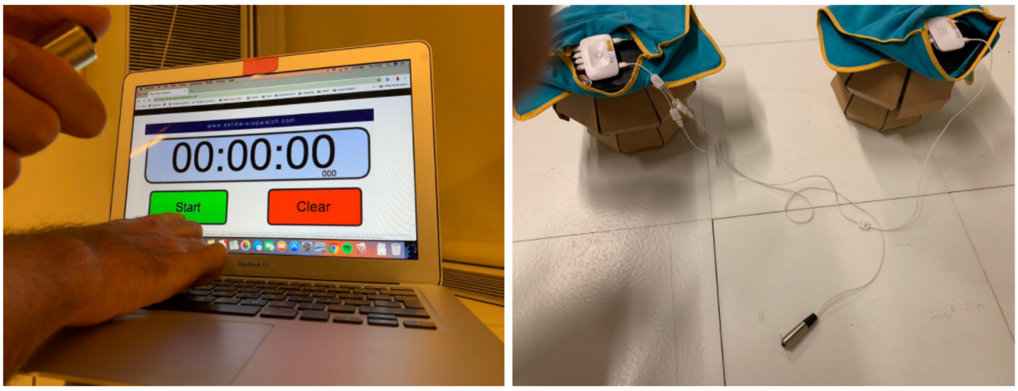


Figure 3.17: The biosignalsplux sync-cables and the multi-sync splitter cable connected to a trigger push button to start offline synchronized acquisition(right) / The universal clock and the acquisition of physiological signals started at the same time (left).

As a next step, we synchronized the tablet questionnaires with the universal clock and noted down the start time of these questionnaires to our experimenter diary. As all the types of questionnaires presented during the experimental sessions were presented through this tablet interface, we were able to see the timestamp of each question being answered. Finally, in order to synchronize the video recordings with the system logs from the Lands of Fog system, we added a hotkey that would be recognized by the system. When the game began, we pressed the hotkey and this action was timestamped in the system log file. This action simultaneously generated an audible notification in the form of a "beep tone" which created a digital record of cut points in video recordings for the start of session. This moment was also synchronized with the corresponding universal clock time and noted in the experimenter diary. This process was also repeated for the LEGO condition.

3.4 Experimental Design

Our experimental design was based on a repeated measures design with two conditions, Lands of Fog environment and A traditional play therapy: LEGO. In this context, we developed the user trials based on dyads of children: a child with ASC (high-functioning) and a child without ASC. The configuration of the experimental setup can be seen in Figure 3.18 in which LOF session is taking place. During LEGO session Lego table moved into the playing area.

Figure 3.18: The configuration of the experimental setup.

3.4.1 Demography

In this research, in total, seventy-two children (36 ASC/non-ASC dyads) participated from the city of Barcelona, with ages between 8-12 years old (N = 12 female, N = 60 male). With the help of our collaborator hospital, the Hospital Sant Joan de Déu, we recruited ASC children who had been diagnosed with high-functioning ASC through the scale of Observation for the Diagnosis of Autism (ADOS) module 3 having a minimum severity diagnosis of 4 [Lord et al., 1999].Module 3 is the grade given for verbally fluent children, while Module 4 is for verbally fluent adolescents or adults. Each module produces a score, which defines the level of autism severity, where 4 is the cutoff value for an ASC diagnosis.

As explained earlier on in this document, we adhere to the vision of people with ASC, which describes that there exists no typical human being since everyone has one condition or another. Because of this, we do not refer to children with ASC and typical children; rather, we refer to children with ASC and children without ASC. This means that we do not control for all the potential conditions that non-ASC children may have (e.g. dyslexia, hyperactivity, ADHD, etc.) and we only control whether they have not been diagnosed with ASC. In this manner we are defining the average population that children with ASC may encounter in their daily activity and hence define an ecologically valid demography for our study. Given this, the children without a diagnosis for ASC that participated in our study, were recruited through dissemination on social media and in schools.

Additionally, both ASC and non-ASC children had to score a minimum IQ level of 70, as determined by the Wechsler Intelligence Scale for Children (WISC) [Wechsler, 1949]. People scoring below 70 are considered to have mental retardation [Gordon and Fleisher, 2010]. It was decided that these exclusion/inclusion criteria mentioned above would allow accomplishing the minimum degree of collaboration needed to go through the game play specified in our experimental conditions, so the child with ASC and non-ASC could play without the assistance of a therapist or parent.

3.4.2 Ethical Approval

All procedures performed were aligned with the 1964 Helsinki declaration, and had ethical approval granted by the ethics committees of the University Pompeu Fabra and the Hospital Sant Joan de Déu (Appendix A.4). This approval process was based on the ethical considerations raised by the research activity with regards to (1) type of data recorded (2) Special needs/Underaged participants (3) mechanisms to protect, anonymize data and process the data (4) Period of data use, deletion of records, transfer of data to others.

3.4.2.1 Type of data recorded

As mentioned in section 3.3 Data Acquisition the type of data recorded included. (i) psychophysiological and body cue data (ii) system data, and (iii) behavior data measured through videos and questionnaires.

3.4.2.2 Special needs / Underaged participants

As the project is aimed towards improving intervention measures for children with ASC, it was necessary to gather data from underaged and special needs participants. We consider underaged participants to have 14 years of age or less, according to Spanish law. As our target range for users was between 8-12 years of age, it can be assumed that all participants were underage at the time of testing. To protect the rights of these minors, all participants were to be only able to participate once written consent is obtained via the informed consent form through their legal tutors and through oral consent from the children themselves. This informed consent form outlines the rights of users within the experimental protocol, aiming for a clear understanding of the facts, implications, and rights of participating in the study, including the right to abandon the experiment at any point. In addition, the legal representative of each participant and a psychologist specialized in children with ASC was present throughout the full experimental procedure, to communicate and manage tasks directly related to the special needs participants.

3.4.2.3 Mechanisms to protect, anonymize data and process the data

All data gathered from the participants in the study is dissociated from the identities of the participants as per the “Instrucció del gerent de 19 de gener de 2012” from Universitat Pompeu Fabra and the “Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995”. Personal data from the subjects were collected by the research team. The team is committed to treat data confidentially and to prevent their alteration, loss, unauthorized processing or access in agreement with the applicable European and Spanish law. Each participant’s ID related information and other information types (physiological data, system data, behavioral data) were associated with a unique alphanumeric ID. To protect the participant’s privacy, the alphanumeric ID file which is associated with participants’ IDs are stored separately from the rest of the documentation in a password-protected account. This alphanumeric ID file is kept on a password-protected disk (data will not be stored on a memory stick or other easily lost/accessed media) which is in a locked cabinet in a closed room, protected with a key. Only the researchers of the project have an access them, under the supervision of the principal researcher. The associated other information types (physiological data, system data, behavioral data, etc.) are kept anonymously in UPF Server which is a server protected by high security measures as established in the Spanish legislation on data protection.

3.4.2.4 Period of data use, deletion of records, transfer of data to others

Within the frame of the funding program of this project, the “Unidades de Excelencia “María de Maeztu””, the data gathered in this project, which is anonymized, is placed in a database (see section 4.3.2 Data sharing) which is open to the research community at an international level and for an indefinite period of time. The intention is that this database keeps growing in the future to improve the knowledge on Autism and on how ICT may have beneficial impact on the social status of children with ASC. The data may be useful for all the ASC research and social community. In this regard, the present set of data gathered will not be able to fulfill some of the articles of the aforementioned directives in the way they are stated due to this extended duration of data and its open access nature. At any time, the participant has the right to request to the researcher access, edit and delete his/her personal data, in agreement with Real Decreto 1720/2007, Ley Orgánica de Protección de Datos de Carácter Personal 15/1999 and EU directive 2002/58/CE. It is again important to mention that the data that is accessible to researchers outside this project guarantees the anonymity and privacy of the subjects and in no case it is possible to reassociate the data to the original participants. For example, in the case of video recordings, only the video coding of the behaviors of the users is available. The video clips are kept under safe storage in our servers and premises as other sensitive information. In case of doing research in body cue recognition, the videos being shared with the research community include face blur filter which does not allow reidentification of the participants.

3.4.3 Procedure

At the beginning of the session, the psychologist welcomed the children and created the first contact with them. In a relaxed environment, the psychologist explained the session and used the visual support tool called "Jumby is Calm" [Gallo and Annenberg, 2009], as used in social skills therapy, so that the children can have an idea of what is coming next and do not generate anxiety linked to the unknown activities they will encounter (see Appendix A.2). Having gained the children's attention and trust by the psychologist, a researcher fitted the child-friendly wearable, designed by our lab for recording the psychophysiological measurements as described in section 3.3.1 Wearable Multichannel Psychophysiology. We used Open Signals software to verify the correct placement of the electrodes (EDA & ECG) and be ready for acquisition to start.

The ASC and non-ASC dyads were exposed to two conditions in a single experimental session: Lands of Fog condition (experimental condition) and the LEGO play activity (control condition). Each pair of children played each of the conditions once for 15 minutes, with a 5 minutes break between them. Before and after each experimental condition, there were 1 min baseline recordings, in a standing position, participants looking at a black screen. Before the experimental conditions and at the end of the experimental session the psychologist applied relaxation exercises (see section 3.3.5 Relaxation Exercises). Before and after each experimental condition, we also administered questionnaires. The assignment of the order of the conditions was carried out randomly to counterbalance the order effect of the tasks. Moreover, before the children played in the Lands of Fog condition, they were introduced to the butterfly nets (see section 3.2.1 Lands of Fog and Figure 3.2) with which they interact in the VE, and watched an introductory video of how they should use it. The goal of this was to avoid the children being frustrated by an inadequate use of the device, as well as to avoid a misuse which could damage the devices. The total duration of the experimental session was approximately 1 hour and 15 minutes (Figure 3.19).

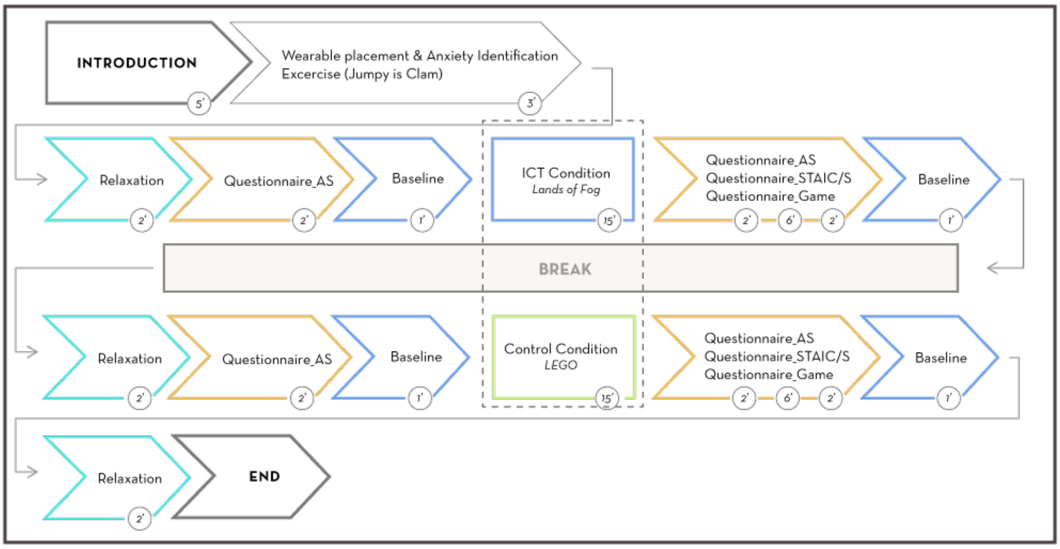


Figure 3.19: Diagram of the Experimental Procedure

Chapter 4

DATA ANALYSIS

4.1 Introduction to this chapter

This chapter introduces the data annotation, preprocessing and feature extraction stage for all the data sources collected in this work. This includes the details of the semi-automatic software pipeline developed for data extraction and the creation of a multimodal database from all these data sources.

4.2 Preprocessing & Feature Extraction

4.2.1 Video Coding

[Parts based on the article *A Structured Video Coding Protocol for Reliable Timestamped Coding of Social behavior for Integration in Multimodal Evaluation: A Case in Mixed Reality for Autism* (under preparation)]

As mentioned in section 2.6.1 Video coding, when coding behaviors during interventions using an interactive system, much richer results can be achieved if the data from video coding is integrated with other data; e.g. system data (log files of system decisions and actions) or psychophysiological data. These are all time series data that keep track of what is going on with the participants at every instant. Relating this data would allow the generation of a shared multimodal representation that could be used in understanding the deeper details of specific behaviors from multiple points of view. This implies that video coding strategies applied on such interactive systems should be flexible enough (on a microanalytic basis) to allow for their integration with different modalities. In this regard, as we saw in section 2.6.1 Video coding, no existing observation scheme was found that allowed reliable time-stamped event coding. We did not find any either that accounted for full-body interactive settings where the system plays a role in fostering SIBs towards or between the children with ASC. Therefore, we developed an adapted observation protocol through an iterative development process.

In particular, we have developed this protocol in the context of a mixed reality experience that fosters SIBs on children with ASC and non-ASC playing together. On the one hand, our protocol proposes a coding grid that has been evolved to successfully code SIBs through a number of projects and iterations. On the other hand, our protocol proposes a systematic step-by-step scheme to video coding that includes a timestamped set of behavior data that can then be correlated with other time series data sources. Results are presented regarding both the tool's flexibility in coding the range of behaviors presented and its ability to achieve good inter-rater agreement scores. This has been done over the course of 20 coding exercises, over a period of approximately one year and a half.

4.2.1.1 Development of Coding Scheme

The social intervention coding scheme defined by Bauminger [Bauminger, 2002] forms the basis for most of the coding schemes that are currently being used for coding social interaction in ASC. This is why we have taken it as the foundations of the coding scheme that we have created. Bauminger's coding scheme is adjusted for interventions for children with ASC from a psychological stance and has already been in use in social skills therapy in another research project by our collaborators from Hospital Sant Joan de Déu. However, it should be noted that this scheme is based on a relatively macroanalytic approach. In their study, [Bauminger, 2002] used this scheme to code child's behaviors in-vivo in 60 seconds intervals during 15 minutes interaction sessions. Specifically, the observer watches the child's behaviors for 50 seconds and then codes them for 10 seconds. Moreover, their inter-rater agreement scores for validating their scheme were based on frequency based comparison (an approach which is used when timing and duration of events are not important in a study) rather than frequency/sequence based comparison (an approach which is used for a detailed indicator of the correspondence between two observations when timing of the events is important) [Illemsen et al., 2009]. Therefore, we adapted Bauminger's method to be able to code in a microanalytic way in both passive and active intervention practices and tested its validity based on a frequency/sequence based approach. We also paid particular attention to the addressee of the social behaviors (i.e. to whom these behaviors were directed) and made necessary adjustments in the Bauminger's scheme which would fit our context. For example, the coding scheme we developed allowed us to differentiate between social interactions which were directed towards the play partner, towards the system, or towards other people who were present in the room but not part of the interaction. This distillation was essential in understanding the effectiveness of an interactive system in promoting SIBs between a ASC child and a non-ASC child while playing face-to-face.

We worked with the BORIS open-source video coding software [Friard and Gamba, 2016], which supports simultaneous video playing from several video sources. Therefore, it allowed us to simultaneously observe the experimental session from the two points of view recorded by the two cameras and microphones placed at opposite ends of the play arena. A native Catalan and Spanish speaker transcribed the children's dialogues from each video to an excel file after the sessions, so that the coders could have a unified version to use when coding the SIBs.

Bauminger's coding scheme is focused on coding social initiation and responses in three sub behavioral categories: positive social interaction, negative social interaction, and low-level interaction. Each category includes a broader variety of prosocial and aggressive behaviors appropriate to the social abilities of high functioning children with ASC. When we reviewed these categories, we agreed that we could have additional items to these category items in the context of our system and could even combine some of the existing category items, such as looking away and avoidance. Furthermore, in Bauminger's study, as we could not find a hierarchical definition for coding social interaction behaviors in a microanalytic way, we decided to develop a hierarchy of interactions and arranged their application order in a flowchart to aid the coders to have a clear step-by-step set of criteria to code. This also required us to define additional categories to detail the social interaction behaviors with categories such as the direction of the social interaction and with the verbal/nonverbal nature of the interaction. Finally, as we gather psychophysiological data, we deemed important to investigate the connection between anxiety related overt behaviors and internal states of the participants at that specific moment. In this regard, we added a category of anxiety indicators which would guide us in detecting such anxiety related behaviors in children with ASC. These anxiety indicators were derived from the Liebowitz Social Anxiety Scale [Liebowitz, 1987] and included behaviors such as physical complaint or intention to end activity. The final version of the coding scheme can be found in Figure 4.1.

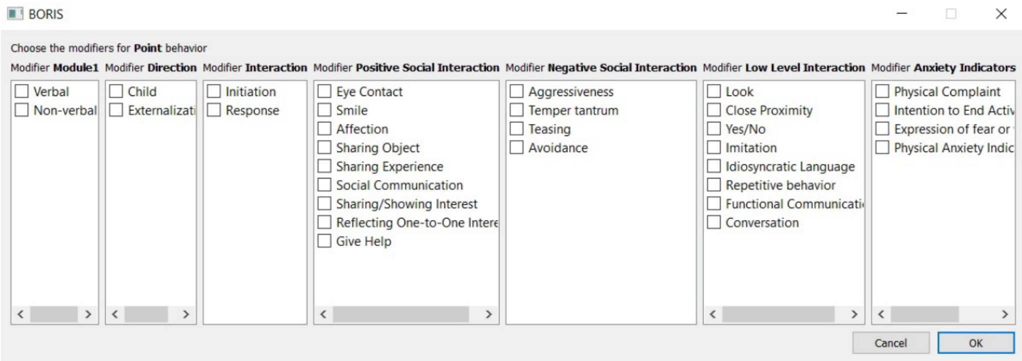


Figure 4.1: Coding categories organized in Boris software format.

In our coding scheme, in parallel with the flow chart, if the behavior detected is a potential social behavior, that behavior can be coded first by selecting one item from each of the first three categories (verbal/nonverbal, direction, and type of interaction), and later selecting items within only one of the following three categories (positive social interaction, negative social interaction, low level interaction). In order to code any anxiety related overt behavior, the coder selects items only from the anxiety indicator category. This approach is explained in four stages in the flowchart (Figure 4.2)

Stage 1: When we began working on a coding scheme for coding SIBs, we followed the following steps:

- We first identified whether a behavior has a potential social value or is an anxiety related behavior.

- After deciding that the behavior has a potential social value, we started interpreting the verbal and nonverbal nature of it.

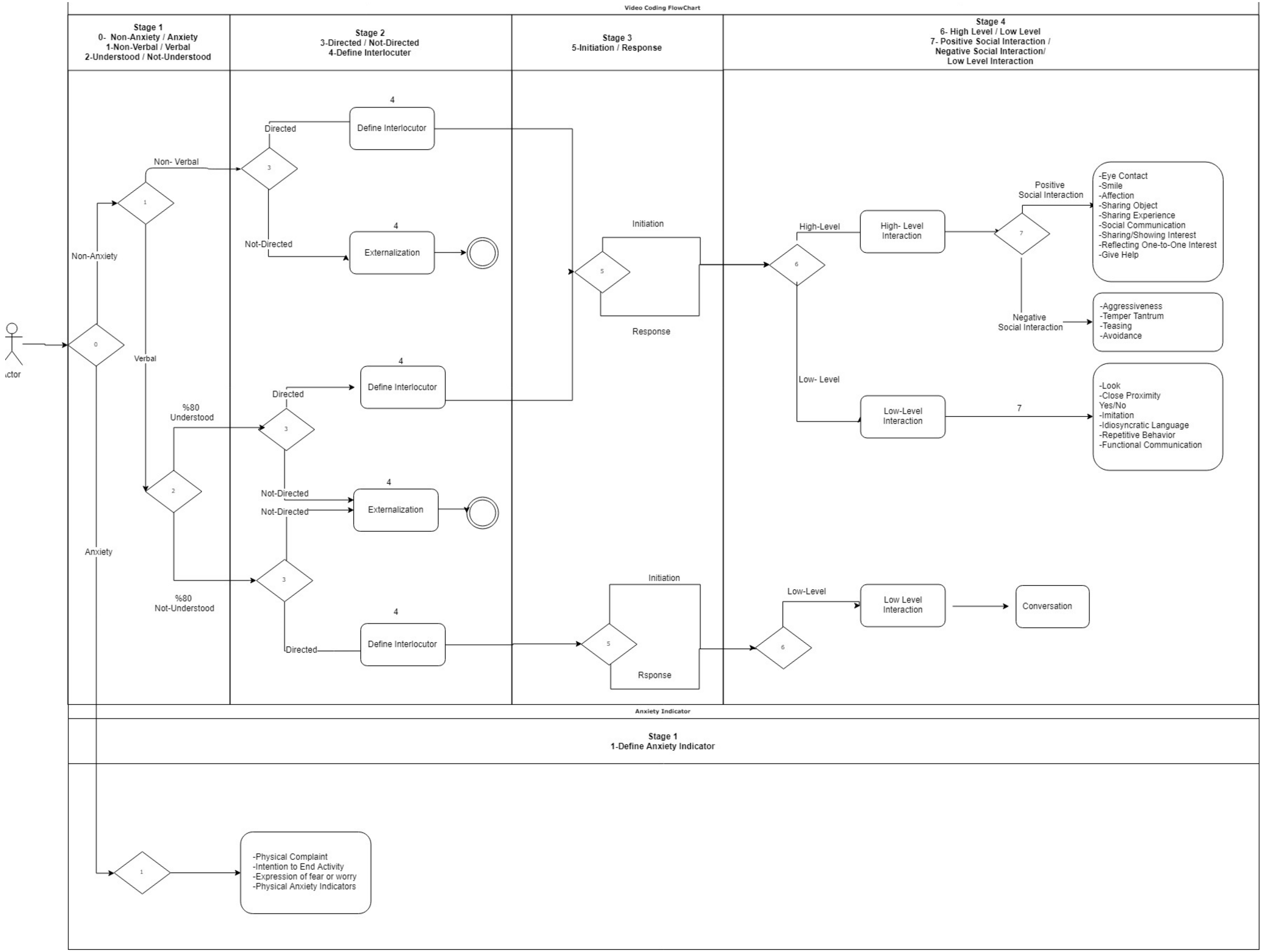


Figure 4.2: Video coding flow chart describing four stages of coding process

- If the behavior is verbal, first we had to make sure that we understood what was being said in that interaction to further our interpretation. For example, it was important to know if comments were appropriate and on subject, or repetitive speech. To facilitate this to coders (as said above) we applied an initial preprocess where a native Catalan/Spanish speaker transcribed the verbalizations to a reference spreadsheet. On the other hand, even if the transcriber could not understand the verbal interaction clearly, we could still use clues on the direction of this interaction based on contextual information. Nonetheless, we provided a differentiation option between clear and unclear verbal interactions to be able to have better sub behavioral categorization in the next stages.

- If the behavior is nonverbal, it was possible to proceed sub-behavioral categorization in the next stages

Stage 2: At the core of the scheme, each social behavior is determined to be an initiation, a response or externalization:

- (1) Initiation: The ASC child begins a new social sequence directed towards the other participant. It must be distinguished from a continuation of a previous sequence by a change in activity;

- (2) Response: The ASC child responds verbally and/or nonverbally to social stimuli directed toward him/her by the other participant;

- (3) Externalization: When the behavior of the ASC child was not clearly directed to anyone in particular, or was addressed at a game element, it was coded as an "externalization" (e.g., self-talk, shouting, dancing). Externalizations are one important adaptation that we have introduced in the coding process. We believe they are important and relevant because, despite them not being formal initiations addressed to the non-ASC child, they do have the potential of calling the peer's attention, as if the behavior was actually addressed to him/her, and hence spark a response.

Stage 3: If the social behavior was directed toward the peer, we then would have sufficient information to code it as an initiation or response.

Stage 4: If the directed initiation/response was clearly understandable, we then would have enough information to mark it as a high or low level interaction. If the directed interaction was not sufficiently clear we coded it as conversation.

We found that certain behaviors began as an interaction at a low level and turned into an interaction at a high level. For example, it was possible to see that a child looks at the peer, then come close to him/her, and finally share an experience. In that case, if they were involved in the same social sequence, we took into consideration only the high level interaction, which we found to be the most important. This sometimes involved watching the sequence until the end and then deciding which aspect of it was the interaction at the highest level. It is also worth noting that each initiation or response needed additional criteria for evaluation of positive or negative, low level behavior similar to the approach being followed by Bauminger.These three behavioral categories were adapted to suit high-functioning children with ASC's prosocial or aggressive behaviors. Externalizations were not associated with these additional criteria since they were not directly associated to a specific social behavior.

All of the coded behaviors were then exported to a summary report in the format of a spreadsheet. However, this was not useful in our case to get a detailed vision on the sequential relationships of the events. Therefore, interactions were converted into visual representation of a series of events where arrows indicate the flow and sequence of events. With its timestamp, the child who started or maintained the interaction is at one end of the line, and the initiation or response qualifier is in the center of the arrow (see Figure 4.3). This was especially helpful when addressing specific social interactions, since diagrams could be aligned to see inconsistencies between coders.

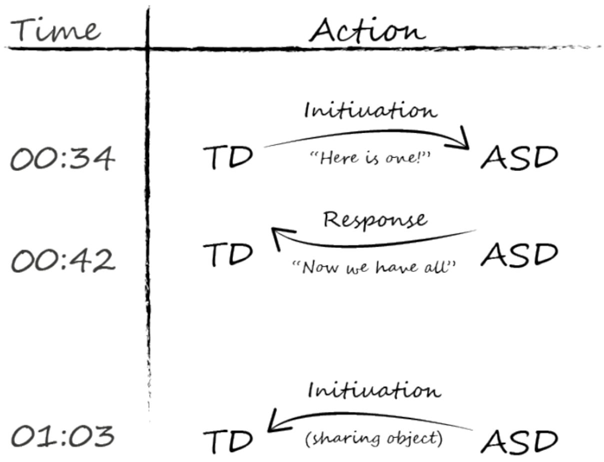


Figure 4.3: Visual representation to understand direction of behaviors

In the case of Bauminger's experiments, as described before, the sessions were coded in vivo, where the children were observed for 50 seconds by a coder and behaviors were coded for 10 seconds. Therefore, the observations were marked for a relatively long aggregation window. The addition of system data and psychophysiological measures to the other types of data collected from users, meant that the rest of data should also be time series data (especially in machine learning approaches that quantify the predictability of a given action during social interactions). Therefore video coding needed to be timestamped to allow a proper multimodal merging of all data, and hence we needed a high degree of time precision in the video coding process. It was therefore natural to use a coding based on a microanalytic method, where precise timing of the events is essential. Also, because of this we did not choose to code in vivo, but rather video recorded the sessions from two side-views, as described above, and then analysed the footage offline, so as to unravel the details of the events and to be able to accurately timestamp the observed behavior.

In adapting Bauminger's coding scheme, we also discussed how to differentiate the strategies between interval based coding (state event) and the one which records every instance of behavior (point event). Aligned with the suggestions from Hauck's study in coding social behaviors [Hauck et al., 1995], we defined the point event as the event which has a temporal window not larger than five seconds. On the other hand, we defined the state event as the event which has a temporal window larger than five seconds ending with a change of theme or action, change of interlocutor, or a response that included a new social initiation from the other child, therefore effectively changing the conversation to a new topic.

For the events which are taking place within the same 5 seconds window, we decided to code only the events which are within the high level interaction category as those events could lead to an effective social process with peers. The exception to this was when a response included concurrent initiation, leading to a conversation on a new topic. As they are both the highest level of interaction we recorded both of the events.

We used an iterative approach in developing the coding scheme and reached a shared understanding after 20 iterations . Each iteration was based on randomly chosen five minutes video segments, from either the first or the last five minutes of both experimental and control conditions. In order to evaluate the validity of the shared conception, in each iteration, we calculated the inter-coder reliability between three coders via Cohen's Kappa using the main social interaction components of the observation scheme (initiation, response and externalization). We used two approaches to compare the coding of these raters, namely "Frequency-based" comparison and "Frequency/Sequence-based" comparison. In our study, as we wanted a high degree of time accuracy in the coded data, the findings from the second approach were critical. Although initially the Kappa scores of the Frequency/Sequence-based method were lower than the results of the Frequency-based method, we realized that there was an increasing trend in the Kappa scores of the Frequency/Sequence-based method at each iteration of the coding scheme. We decided to finalise the iterative refinement process of our protocol when we reached a level of agreement of 0.68-08.

4.2.1.2 Validation of the Coding Scheme

The validation of the coding scheme consists of two stages. In the first stage, we investigated the inter-rater reliability with a selection of data and in the second stage we assessed the applicability of the scheme by comparing the data from all coded videos. The coded videos were from 18 experimental trials (see section 3.4 Experimental Design). The set of videos was composed of the first and last 5 minutes of each play session(resulting in coding 72 video segments of 5 minutes).

4.2.1.2.1 Inter-rater Reliability

The inter-rater reliability of the video coding scheme was calculated for each category (occurrence of social initiations, responses and externalizations), both through percentage agreement and Cohen's Kappa. Kappa scores were between 0.60 to 0.69 and the percentage level of agreements were between 0.71 to 0.78 with three coders, providing reliability of the data.

4.2.1.2.2 Comparing Play Behaviors

The descriptive information obtained when coding Lands of Fog and LEGO conditions is described in the findings below. These findings present the coding scheme's flexibility in coding the range of behaviors presented in a technological intervention and a non-technological intervention.

Overall, the most common high level behaviors coded were "Sharing Experience" and "Sharing/Showing Interest" in both conditions. While coding "Sharing Experience" we adopted the same definitions presented in Bauminger's coding scheme "The child tells about an experience to peers or asks them about their experiences (e.g., "What did you do over the weekend?")". On the other hand, "Sharing/Showing Interest" was defined mostly in the context of pointing action which might be detailed according to the intention of the person, as well as by the linguistic function it serves (e.g child says "look at this" to his play partner and/or uses pointing gesture). The most common low level behavior in the MR condition was "Close Proximity" (defined as standing in close proximity to another child (3 feet or less) but does not approach the peer), and the most common low level behavior in the LEGO condition was "Look" (defined as looking at the other child's face or body, or child's action, without establishing eye contact).

Of the behaviors coded in MR condition, 71.0 percent of behaviors were clearly directed at the play partner, while 29.0 percent of behaviors were considered "externalizations". Likewise, in the LEGO condition, 72.8 percent of behaviors were directed at the play partner, while 27.2 percent were considered "externalizations". Of the directed behaviors in MR condition, in terms of initiations, 80.3 percent of them were verbal and only 19.7 percent were nonverbal. For responses, 67.8 percent of them were verbal and 32.2 percent were nonverbal. In the LEGO condition, a similar trend was recorded. Finally, in the MR condition with 64.3 percent of initiations indicated as high level and 35.7 percent low level, we could see that initiations were more likely to be high level behaviors than low level behaviors. For responses, high level and low level behaviors were evenly split at 50 percent each. A similar trend was observed in the LEGO condition.

4.2.1.3 Applicability

The purpose of this study was to develop and validate a protocol for coding SIBs for ASC in technology-based social intervention systems, designed to achieve reliable timestamped coding for integration in multimodal time series evaluation.

During the process of developing the coding scheme, the team frequently did coding checks, applied an iterative development approach, and made necessary adjustments to reach a shared understanding. This process helped us to find uncertainties in the current coding protocol and adjust it to a version that is suitable for our context. In order to achieve acceptable percent agreement scores, an extended development period was required before having the protocol sufficiently clear and precise so that the coders can get trained in a reasonable time and understand it unambiguously so that they all code the same.

When our final coding protocol was applied to a randomly chosen new video, the calculated reliability scores were found acceptable according to current standards set by the WWC [Clearinghouse, 2013]; where agreement levels are deemed acceptable at or above 0.60 for Kappa and 0.80 for percentage agreement. Furthermore, since in both conditions, the proportion of coded SIBs (and the associated subcategories) is similar, it can be inferred that the developed coding technique is flexible enough to address the coding requirements of both conditions.

As we aimed to develop a standardizable coding protocol which could be used across different settings, the distillation of direction of behavior helped us to filter out the behaviors directed toward the game activities. In other words, if our results were solely based on this coding scheme's effectiveness in finding game directed social behaviors, this would be a questionable finding in terms of its applicability to different settings. This separation was especially challenging in the initial iterations of the development where we had also included coding behaviors directed towards game events such as the moments the ASC child talks to his character. Such acts were confused with the SIBs we wanted to foster between peers during the coding process. In this regard, a consensus was reached that interactions directed towards the game in any capacity should not be coded as SIBs in the context of this project, as the purpose of the game and characters was to foster SIBs between the child with ASC and the non-ASC peer. However, another version of this coding protocol might incorporate new categories for SIBs directed towards game characters, researchers, and other individuals present.

4.2.2 EDA signal processing and feature extraction

The EDA signal is made up of the superposition of two signals: the tonic level of skin conductance (SCL), representing the baseline signal, and the superimposed phasic increases in conductance. The phasic components reflect a unitary skin response (SCR). In turn, the responses are given by the activity of the eccrine sweat glands in response to external stimuli [Fowles et al., 1981]. The decomposition of both components of the EDA signal was done through a software package for MATLAB called Ledalab 3.4.9 [Benedek and Kaernbach, 2010]. The Automatic EDA Artifacts Identification library EDA Explorer [Taylor, S., and Jaques,] was used to identify possible artifacts in the samples and then, they were corrected by interpolation. The movements of the body did not generate strong artifacts, so the features of the signal could be correctly extracted. Subsequently, the deconvoluted signal was analyzed by the default peak detection algorithm. To detect significant peaks, the local maximum must have a difference greater than 0.01 microsiemens compared to the previous peak or must follow a local minimum [Braithwaite et al., 2013]. The phasic features were calculated within different response windows based on the characteristics of the interaction under investigation. For example, when we wanted to investigate the ANS activation related to each event (such as social initiation moment or character merge) taking place during experimental conditions, the phasic features were calculated within a response window (rw) up to 4 seconds in length following the event onset which is a commonly used approach [Posada-Quintero and Chon, 2020]. When we wanted to investigate the ANS activation related to the whole experimental condition, the phasic features were calculated within a response window (rw) up to 15 min in length following the start of each experimental condition. All the features included in the Ledalab's analysis spectrum were extracted [Benedek and Kaernbach, 2010]. We used different normalization techniques for each feature extracted including using z-scale transformation and subtracting the corresponding baseline values from experimental condition values.

4.2.3 ECG signal processing and feature extraction

The HRV indicates the parasympathetic nervous system (PNS) index, which is of great interest since greater PNS activity of the vagal tone is associated with better social functioning [Laborde et al., 2017]. HRV was computed from the raw ECG signal obtained through the electrodes of the wearable; and the signal data was imported into Kubios [Tarvainen et al., 2014]. The samples were analysed manually for possible artifacts. The processing of the signal features was performed again in different time windows based on the characteristics of the interaction under investigation. For example, when we wanted to investigate the HRV related to each event taking place during experimental conditions, the features were calculated within a response window (rw) up to 30 seconds in length following the event onset which is minimum window size allowed to be used Kubious and common approach for event based HRV analysis [Tarvainen et al., 2014]. When we wanted to investigate the HRV related to the whole experimental condition, the features were calculated within a time window up to 15 min in length following the start of each experimental condition. In these windows there were series of R-R intervals which refers to the time between two R peaks of a traditional ECG signal. All features were extracted in the time-domain, the frequency-domain, and the nonlinear indices that the spectrum of Kubios includes. We used different normalization techniques for each feature extracted including using z-scale transformation and subtracting the corresponding baseline values from experimental condition values.

4.2.4 ACC signal processing and feature extraction

The biosignalsplux's Accelerometer is tri-axial and measures sub-milliG accelerations and provides raw data of each axis on individual channels. Although the primary purpose to use ACC in this research was to detect movement artifacts in psychophysiological measures, we also investigated the potential of this raw signal in detecting physical activity or activity count (AC) in experimental conditions. However, there is a rich variety of approaches used for data collection and analysis because of the increasing availability of ACC-based sensors in physical activity detection, and there is currently no widespread scientific agreement on which methods to use.Therefore the correct decisions on which parameters should be employed in a specific case are challenging for the researchers and practitioners. Moreover, up to now, far too little attention has been paid to developing an explicit, open-source, and reproducible summary metric based on raw data [Migueles et al., 2017]. These steps are essential for research into physical activity, as current definitions of AC include patented and device/software specific definitions [Bai et al., 2016].

4.2.5 Body Cues

[Parts based on the article *Sayis, B., Pares, N., Gunes, H. (2020, October). Bodily Expression of Social Initiation behaviors in ASC and non-ASC children: Mixed Reality vs. LEGO Game Play. In Companion Publication of the 2020 International Conference on Multimodal Interaction (pp. 140-149).*]

In our body cue analysis our focus was on the moments of SIBs. In this regard, to include the complex dynamics of body gestures in our analysis, we defined clips of the SIBs extracted from the video coding process as temporal windows of 10 seconds (e.g 5 seconds before and 5 seconds after the coded social initiation moment).

4.2.5.1 Automatic body and head pose tracking

To obtain the full body skeleton joints, we used the fully automatic multi-person 2D body tracker OpenPose [Cao et al., 2018]; which provides the joint locations of the human body (e.g., hand, shoulder, etc.) in RGB sequences. Similar to the approach followed in [Coppola et al., 2016], we focused on 15 joints for extracting features related to body pose. To track head pose we adopted the Openheadpose framework [Cao and Can,] which leverages on OpenPose. In order to characterize the labeled social interaction clips, we extracted body features using tracked skeleton information. These features were divided into two main categories: (i) individual features with n=51 for each individual totaling to 102 individual features; and (ii) interpersonal features with "proximity" (prox) n=245, and "visual focus of attention"(vfoa) n=6, in accordance with the labeled social initiation.

4.2.5.2 Spatio-Temporal Features from Individual Data

In order to obtain features from both subjects, we implement the feature extraction methodology of [Faria et al., 2014] and [Faria et al., 2015] which have been used successfully in human day-to-day activity detection. These features also encapsulate majority of the features that Kleinsmith [Kleinsmith et al., 2005] first introduced in 2005 where they have attempted to understand whether emotion categories and affective dimensions can be predicted from affective postures. Following [Faria et al., 2014] and [Faria et al., 2015], we obtained 51 spatio-temporal features, such as: euclidean distances between joints; angles formed between joints; and torso inclination. Each individuals' individual features grouped under the category individual features with 102 features in total.

4.2.5.3 Spatio-Temporal Features from Interpersonal Data

Gaze is considered the primary focal point of social interactions [Mason et al., 2005]. Visual Focus of Attention (vfoa) is a version of gaze that reveals "who is looking and where" [Sheikhi and Odobez, 2012]. In order to detect the vfoa of each participant we used the head pose estimation features expressed as angles of pitch, roll and yaw for each individual (6 features in total). We also used Hall's [Hall, 1963] proxemics theory, to estimate proximity features. Since the MR system facilitates the collaborative activities on joint actions, children's proximity might be a factor leading to initiations. In this regard, similar to the strategy followed in [Coppola et al., 2016], we define social features as the ones that describe the relationship between two skeletons based on physical proximity, i.e. inter-body distance during the interaction. This set of features encompasses different subsets of features as follows: (a) log-covariance of the joints distances between bodies (120 features); (b) the minimum distance between any joint of one person and the torso of the other (2 features); (c) torso to torso distance (1 feature); (d) the energy over the euclidean distances from all joints of skeleton one to the torso of skeleton two and vice-versa (2 features); and (e) features that were computed similarly to the first subset, however, in a temporal way (a temporal window, herein defined as 10 frames). From all subsets, we extracted 251 interpersonal features per frame given both skeletons are interacting, which is similar to other feature sets. We formed our feature vectors per frame given that each 10 second clip for social initiation moments consists of 300 frames (recording 30 fps). As an initial step, we resampled every social initiation clip and it was done with respect to the mean, computed over 300 frames for each social initiation clip.

4.2.6 System Logs

As the frequency of targeted behaviors elicited during the free play sessions could help us to understand the effectiveness of the interaction design strategies we are using, the total number of occurrences of each event is also extracted as new features in our data analysis. Moreover, we also monitored the activity level of children during the session as a measure of their awareness and interest in engaging in active play. In this regard, we calculated the distance covered by each participant which was calculated based on the trajectory of their nets in the playing plane (area_covered). Moreover, as the distance between participants could give us an insight on their level of comfort when it comes to deciding to play collaboratively instead of playing alone, we calculated the distance between their nets in the playing plane (distance_btw_participants).

4.3 Database Construction

In terms of investigating the relationships between the processed data sources, we created three main data categories. (1) psychophysiological and body cue data (2) system data, and (3) overt behavior data (Figure 4.4).

Possible relationships between these 3 categories were investigated and data was organized based on two different data analysis strategies (1) event-related responses and (2) frequency of responses. In event-related responses, psychophysiological reactions and body cues can be attributed to a specific eliciting stimuli (e.g social initiation moment, character merge, etc) for the specified phase of the experimental condition and for the specified type of participant. On the other hand, in frequency of responses strategy, overall psychophysiological reactions and body cues can be attributed to the rate of specific events and questionnaire reponses (e.g number of social initiation moments, number of character merge events, STAIC questionnaire responses) for the specified phase of the experimental condition and for the specified type of participant.

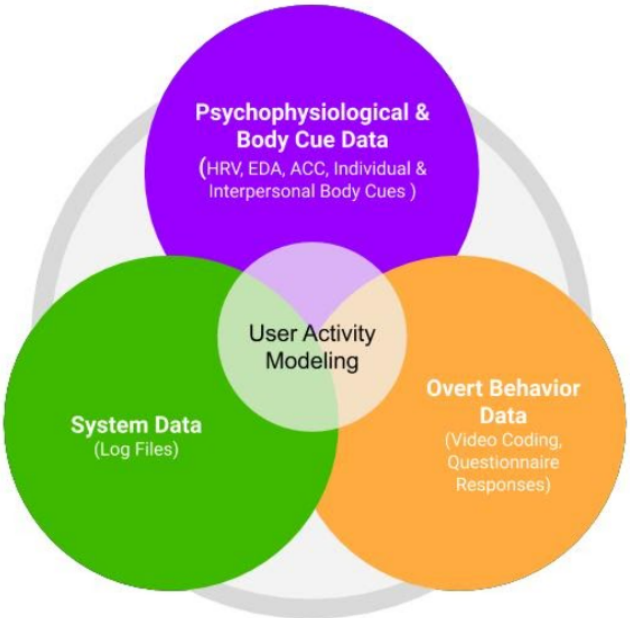


Figure 4.4: Diagram of the three main data categories. (1) psychophysiological and body cue data (2) system data and, (3) overt behavior data.

4.3.1 Semi-automatic data extraction

Based on the strategies mentioned in the previous section we developed a Python-based semi-automatic software pipeline. We developed a scalable data pipeline architecture to train machine learning models with the multiple reconstructed datasets (Figure 4.5). We extracted datasets associated to frequency of responses (Figure 4.6) and event related responses (Figure 4.7).

This software pipeline is based on a software project which consists of two parts. The first part is mainly focusing on the data source separation, data annotation, and data preparation for the use of Ledalab and Kubios preprocessing and feature extraction programs. The second part is focusing on remaining feature extraction processes, data fusion and data frame construction for the event related responses and frequency of responses strategies.

4.3.1.1 Frequency of responses dataframe

The frequency of responses dataframe (Figure 4.6) is grouped by categorical variables, having four categorical variables: "experiment_no", "participant", "condition", "phase" and four different types of quantitative data "Overt behavior", "System Logs", "Psychophysiological Measurements", and "Body Cues", where each one has sublevel features sets associated with the features extracted. For example, the frequency of responses dataframe consists of:

- data from 36 different trials (experiment_no),

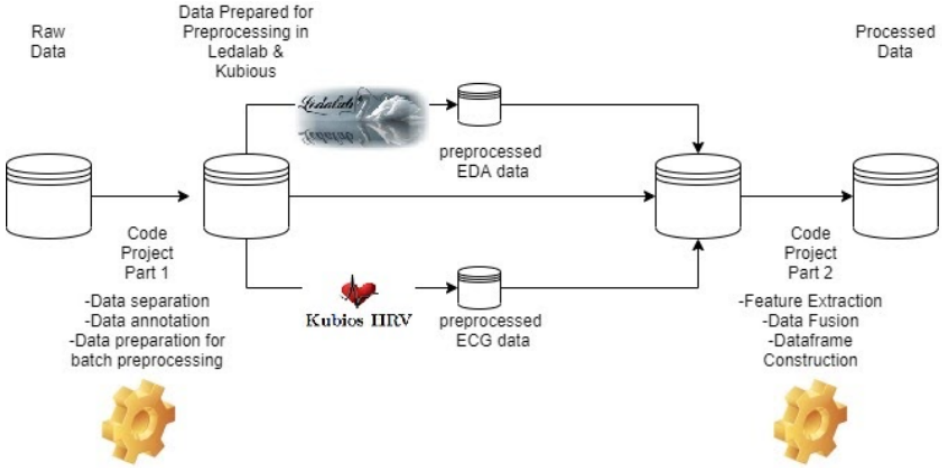


Figure 4.5: Python-based semi-automatic software pipeline for database construction.

- in each trial we have data from ASC and non-ASC participants (participant),
- each participant participated in LEGO and LOF condition (condition),
- each condition data is formed by aggregating interval data from five different phases (phase): data from the first 5 minutes of the session, mid 5, last 5, before and after the session respectively.

This dataframe was organized in a way to give a summary measure over the specified phases based on the quantitative data collected. For example, the individual level data from this aggregated data for Overt behaviors has data from video coding and questionnaires. In video coding we have a feature set of three (n = 3): the number of initiations, number of responses, and number of externalizations. Since video coding features can only be associated with the during the session intervals, we do not have data for pre and post baseline intervals for this feature set (The corresponding cells are highlighted with grey color in Figure 4.6). Moreover, since we only coded first and last five minutes of the sessions we also do not have data for the mid 5 interval for video coding featureset. Likewise, the feature set questionnaires has nine features which can be only associated with the total duration of the sessions. System Logs feature set has 10 features which can be only associated with the LOF condition. Moreover, since currently we have only extracted body cue information for the specific social interaction moments, we do not have aggregated body cue data for the specified session intervals.

4.3.1.2 Event related responses dataframe

The event related responses dataframe (Figure 4.7) is grouped by categorical variables, having six categorical variables: “experiment_no”, “participant”, “condition”, “phase”, “event_type”, “event_id” and two different types of quantitative data “Psychophysiological Measurements”, and “Body Cues” each has sublevel features sets associated with the features extracted. For example, the frequency of responses dataframe consists of

				Overt Behaviour		System Logs	Psychophysiological Measurements		Body Cues	
experiment_no	participant	condition	phase	video coding (n = 3)	questionnaires (n=9)	system logs (n= 10)	EDA (n = 12)	HRV (n = 49)	individual (n = 51)	interpersonal (n = 251)
1...36	ASC / non-ASC	LEGO	first5							
			mid5							
			last5							
			total							
			pre_baseline							
			post_baseline							
		LOF	first5							
			mid5							
			last5							
			total							
			pre_baseline							
			post_baseline							

Figure 4.6: The frequency of responses dataframe.

- data from 36 different trials (experiment_no),
- in each trial we have data from ASC and non-ASC participants (participant),
- each participant participated in LEGO and LOF condition (condition),
- each condition has data from five different phases (phase): data from the first 5 minutes of the session, mid 5, last 5, before and after the session respectively,
- in each phase there are events from social interaction behaviors (initiation, response, externalization) and game activity information (extracted from system log events such as the moment of character_merge),
- each event takes place in different order denoted with an specified id (event_id).
- for each specific event we have information based on psychophysiological and body cue measurements.

This dataframe was organized in a way that each event can be analyzed independently from each other based on psychophysiological and body cue measurements. However, currently body cue features are only extracted from social interaction moments such as initiation, response and externalization not from game events (The corresponding cells are highlighted with grey color in Figure 4.7). Moreover, mid5 of each condition does not currently include video coding events as explained before. Likewise, LEGO condition does not include events associated with the game activity.

						Psychophysiological Measurements		Body Cues	
experiment_no	participant	condition	phase	event_type	event_id (1 ... N)	EDA (n = 12)	HRV (n = 49)	individual (n = 51)	interpersonal (n = 251)
1...36	ASC / non-ASC	LEGO	first5 / mid5 / last5 / total / pre_baseline / post_baseline	video coding - initiation	initiation #1				
					initiation #2				
					.				
					.				
					initiation #N				
				video coding - response	...				
				video coding - externalization	...				
		LOF	first5 / mid5 / last5 / total / pre_baseline / post_baseline	video coding - initiation	...				
				video coding - response	...				
				video coding - externalization	...				
				system log - hunted_insect	...				
				system log - character_created	...				
				system log - character_changed	...				
				system log - character_points_at	...				
				system log - character_being_idle	...				
				system log - manipulate_prop	...				
				system log - character_merge	...				
				system log - character_greeting	...				

Figure 4.7: The event related responses dataframe.

4.3.2 Data sharing

To allow sharing and reuse of the raw data and the data extracted, as a digital repository, we used Zenodo. Zenodo uses DataCite to register DOIs for all submissions and hosts these submissions using the safe and reliable foundation of CERN's data centre. next to the biggest scientific dataset in the world, the Large Hadron Collider (LHC)'s 100PB Big Data store. This means that the code and data in Zenodo will be retained for the lifetime of the repository.

The majority of the data is comma separated variable (CSV) format. Data files, images and videos are included in the open data sets in an anonymized manner (face-blurred), aligned with the consents from the participants. A document record and change track is included in a separate metadata file for all published files.You can access to this dataset with DOI: 10.5281/zenodo.4557383

The code projects related to end to end pipelines in order to preprocess, classify and analyze the raw data from these trials are also organized and uploaded to github open repository (https://github.com/batuhansayis/ASCMEOR.git).

Chapter 5

USER ACTIVITY MODELING

5.1 Introduction to this chapter

In this chapter, we discuss the computational approaches that have been used in this research for modelling SIBs in children with ASC in our MR environment with the experience Lands of Fog, and to compare to the traditional therapy setting. This modelling has been based on psychophysiological data, body cues, self assessment and game activity measurements as described in previous sections. Specifically, we propose three different computational models by following two different strategies. In the first strategy, we focus on investigating the effect generated by each social initiation moment in psychophysiological measurements and body cues separately. In the second strategy, we focus on investigating the observed general social state of the children with ASC in relation to all the data collected during the sessions. We present and discuss the results step by step. These discussions include the ideas presented in the previous chapters, linking concepts from the perspectives of design, implementation, and evaluation of the full-body interactive MR environment for fostering SIBs in children with ASC

5.2 Psychophysiological Modeling of SIBs

[Parts based on the article *Sayis, B., Crowell, C., Benitez, J., Ramirez, R., Pares, N. (2019, September). Computational modeling of psycho-physiological arousal and social initiation of children with autism in interventions through full-body interaction. In 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII) (pp. 573-579). IEEE.*]

5.2.1 Introduction

In this study, our goal was to run an initial investigation with the already available data towards understanding the effectiveness of Lands of Fog when compared to LEGO condition based on the psychophysiological arousal generated in social initiation moments in children with ASC. We used machine learning techniques to analyze the multimodal data from the ASC children obtained during 18 trials (3 female and 15 male). As a result, the proposed computational model which is based on psychophysiological measurements helped us see that certain HRV and EDA features which might be associated with social functioning, anxiety, or physical activity could have an effect on finding the differences between the social initiation taking place in the Lands of Fog condition and that taking place in the LEGO condition.This was the first sign showing that our MR system has specific properties, compared to a traditional construction-based intervention, which potentially provide a new interesting context to intervention in ASC

5.2.2 Data

The data being used in this study was based on the reconstructed version of the "Event related responses dataframe" and included only the data from the first 18 trials and only from the ASC participants in social initiation moments with corresponding psychophysiological measurements.

5.2.3 Experiments & Results

As the degree of alteration in psychophysiological arousal among individuals with ASC during social interaction might correspond to the severity of social difficulties experienced [Neuhaus et al., 2016], we wanted to test this hypothesis in our comparative study focusing on the social initiation moments in which children with ASC are least successful in generating it by themselves.[Sigman et al., 1999]. In this regard, for this problem we defined a 2-class classification task. The initiations taking place in the Lands of Fog (LOF) MR condition were named "ASC LOF initiation" and the initiations occurring in the LEGO construction condition were named "ASC LEGO initiation". The classes consisted of a balanced number of instances (N = 177 ASC LOF initiation, N = 198 ASC LEGO initiation).

We used Weka Data Mining Software for automatic classification [Hall et al., 2009]. We applied modality fusion to integrate all incoming single modalities (EDA and HRV) into a combined single representation. In order to fuse the EDA and HRV information, we implemented feature-level fusion by using the extracted features of each modality and combined these features into a single large vector. The resulting feature set was quite large (50 features, 12 EDA features and 38 HRV features). Therefore, it was mandatory to use a feature selection technique to find the features from both modalities that maximized the performance of the classifier. We used different feature selection strategies for our dataset. We also explored whether the model accuracies could be further enhanced by gathering more data and applying best performing algorithm in different sizes of the same dataset.

5.2.3.1 Feauture Selection

5.2.3.1.1 All Features (FS1): 50 features (EDA 12 features and HRV features 38) were used as a result of feature extraction.

5.2.3.1.2 InfoGain (FS2): Automatic feature selection using filter techniques was used to acquire the most appropriate subset of features for classification. To score the features, filter techniques use a proxy measure (e.g information gain), where these features are ranked by information gain values. The list of the descriptors taken into account for this approach with corresponding features is presented in the InfoGain row of Table 5.1. The improvement in the accuracy in training the models over distinct feature subsets was also investigated, ranging from 1 feature to all features (incremental addition of features based on their information gain rank). As a result, learning curves were generated over the number of instances in order to evaluate the model's performance over incremental feature subsets (as shown in Figure 5.1). Figure 5.1 shows that the addition of each feature did not have a significant impact on the accuracy of the model. Consequently, the entire set of the features described in Table 5.1 was used for the InfoGain based modeling in the following sections.

Table 5.1: The Features Extracted from HRV & EDA Signals

Category	**Features**	**Signal**
All	Time-domain, Frequency-Domain, Nonlinear Features	HRV
	Continuous Decomposition (CDA) & Standard trough-to-peak (TTP) Analysis	EDA
InfoGain	MeanHR(1/min), MeanRR(ms), EDR(Hz),HF(ms^2), SD1(ms), pNN50(%), RMSSD(ms), NN50(count), SD2(ms), STDRR(ms), Totalpower(ms^2), $TINN(ms)$, RRtriindex, LF(Hz),LF(ms^2), $VLF(\%)$, $HF(\%)$	HRV
	GlobalMean	EDA
Manual	MeanHR(1/min), MeanRR(ms), RMSSD(ms), pNN50(%)	HRV
	GlobalMean, CDAAmpSum, CDASCR, CDAISCR	EDA

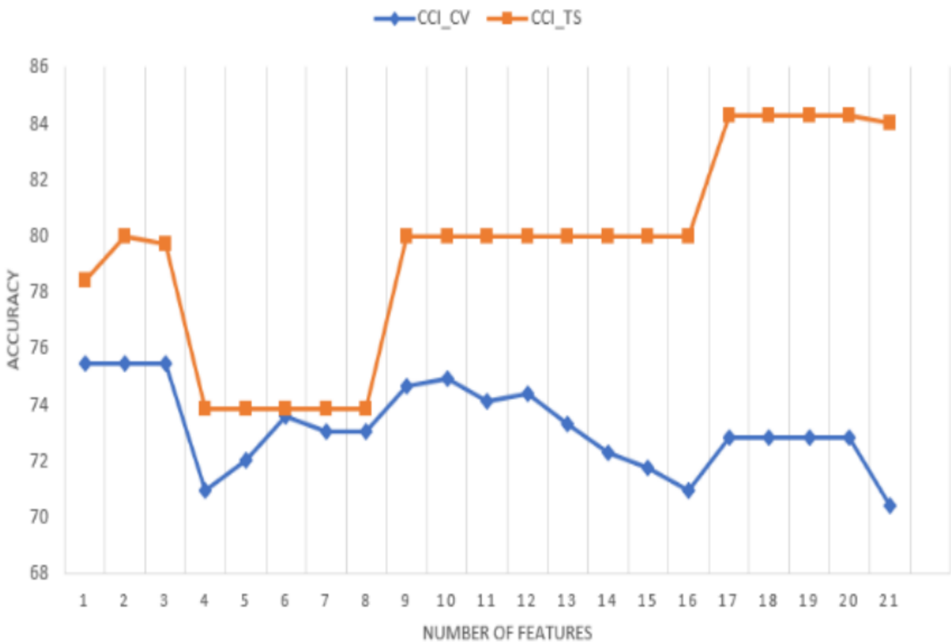


Figure 5.1: Learning Curves showing the increase of the number of features organized by information gain based on comparison of Correctly Classified Instances (CCI) of Cross Validation (CV) Training Set (TS).

5.2.3.1.3 Manual (FS3): Figure 5.1 shows some of the highly ranked HRV features improving the overfitting problem until the eighth iteration. When these features were investigated we saw that the feature set in the eighth iteration also includes some of the features, which best reflect the vagal tone, such as mean square of successive differences (RMSSD), pNN50 (the percentage of successive normal sinus RR intervals more than 50 ms) and HF (high-frequency HRV) [Otzenberger et al., 1998]. As a result we decided to manually create a smaller feature set keeping the two highest ranked features, MeanHR (1 min) (the mean heart rate) and MeanRR(ms) (the mean of Inter-beat-interval) together with RMSSD and pNN50 excluding HF which might not provide a reliable assessment of vagal tone for the rw lower than 1 min [Malik et al., 1996]. A similar strategy was also applied in the manual EDA feature selection step by choosing the EDA features, which represents the phasic activity most accurately [Benedek and Kaernbach, 2010] such as CDASCR (Average phasic driver within rw), CDAAmpSum (Sum of SCR-amplitudes of significant SCRs within rw) and CDAISCR (Area of phasic driver within rw) and creating a feature set with the highest ranking EDA feature GlobalMean (Mean SC value within rw).

5.2.3.2 Algorithms

We compared the results of commonly used algorithms such as Neural Networks (NN) (with one hidden layer), Support Vector Machine (SVM) (with a linear kernel), Decision Trees, K-nearest neighbor (KNN) (Number of nearest neighbors K = 1) and Logistic Regression using Weka Data Mining Software [Hall et al., 2009], utilizing the default parameters.

5.2.3.3 Results

We trained a model for each feature selection strategy in order to investigate the effectiveness of the model. Thus, the results we present include a comparison between the feature selection strategies as well as the algorithms. We evaluated the accuracy as CCI% using 10-fold CV (Table 5.2). The table below shows that the accuracy increased when feature selection strategies were considered (feature selection FS2 and FS3). Significant improvements were achieved by the algorithms NN and Decision Trees, 77.1 being the highest accuracy obrained with the FS3 feature set with the Decision Tree Algorithm. We have also investigated whether the best performing model (Table 5.2) could still be improved by collecting more training data over a varying number of training instances (Figure 5.2).

We created different subsets of training instances. Each subset was organized by randomly removing instances from each trial's data while preserving the same distribution as the full dataset. As shown in Figure 5.2, there was no significant change in the model's accuracy although there was a change in the size of the training sets.

5.2.4 Discussion & Conclusion

In this work, we investigated the psychophysiological arousal activity in our full-body interaction MR experience by means of HRV and EDA recordings during social initiation

Table 5.2: Class Classification Accuracies Measured as CCI% for Training Set (TS) & 10-Fold Cross Validation (CV)

	All (FS1)		**InfoGain (FS2)**		**Manual (FS3)**	
Classifier	TS	CV	TS	CV	TS	CV
Baseline	52.8	52.8	52.8	52.8	52.8	52.8
Decision Trees	86.1	69.3	84	70.4	82.6	**77.1**
ANN	97.6	66.6	90.4	**76**	76.2	71.4
SVM	73.3	**71.5**	70.9	67.7	70.4	69.3
Logistic Regression	75.5	67.2	72	66.1	71.2	68.2
k-NN	85.1	67.3	84.8	72.2	80.8	68.8

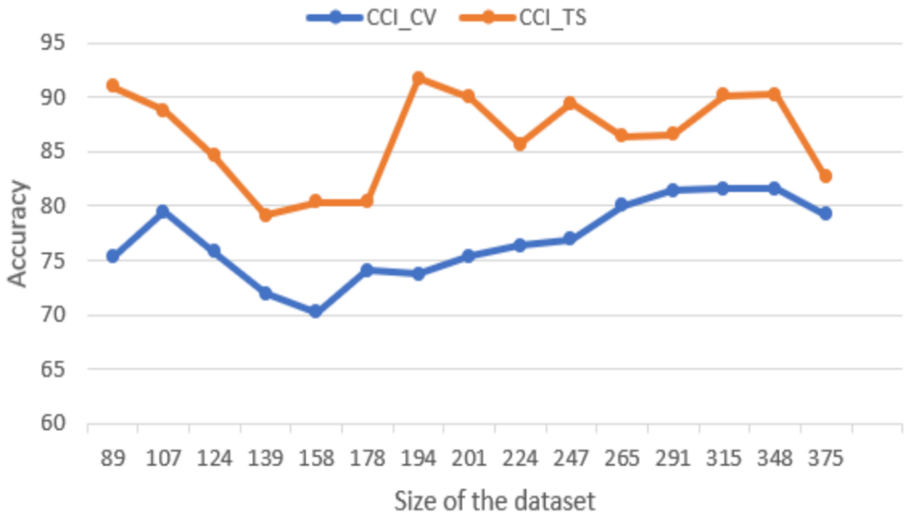


Figure 5.2: Learning Curves showing the increase of training instances.

behaviors in children with ASC playing with a non-ASC child. As the initiation behaviors of ASC children during the Lands of Fog MR experimental condition and LEGO control condition might relate to possible change in the arousal activity, the problem to be solved was defined as a 2-class classification task. For distinguishing between these 2 classes we used the extracted features from HRV and EDA recordings and applied different feature selection strategies which maximize the performance of the classifier. We have found that the classification accuracy increased when feature selection strategies were considered (feature selection strategies FS2 and FS3).

In FS2, the HRV feature set included some of the features which best reflect the vagal tone such as RMSSD and pNN50. On the other hand, the EDA feature set only included the feature GlobalMean while excluding the features that represent the phasic activity most accurately. A possible explanation for this contrast could be that the features selected by FS2 might have been contributing mostly to the classification of physical activity, instead of the psychophysiological arousal activity. Moreover, the GlobalMean feature was the highest-ranking feature among HRV and EDA features. This feature may reflect the thermoregulation based on physical activity [Cacioppo et al., 2007]. Additionally, thermoregulation may play a bigger role in EDA measurement location, when comparing shoulder electrode placement to other parts of the commonly used EDA measurement locations (i.e. finger palm) [Turpin et al., 1983]. This may be supported with our own results regarding our recordings during summer time, where we were initially unable to control the room temperature, leading us to having out of range SCL levels for two of the participants. As a result, we had to exclude these recordings in our analysis.

We are aware that physical activity might have an influence on HRV and EDA because of the reasons mentioned before. Because of this, we designed the two conditions (experimental Lands of Fog experience and control LEGO activity) as similar as possible regarding the ambulatory nature of the activity and the exertion which the users have to go through. Therefore, this placed them in a balanced context so that they could be compared without physical activity getting in the way of our model. Hence, this feature was not found to be relevant. Moreover, when the EDA features selected included the features which are representing the phasic activity most accurately, we saw that the model's accuracy increased. The main phasic activity related EDA features formed in FS3 contributed to this classification to reach higher accuracy rates. Likewise, in FS2 (the HRV infogain feature set) we see MeanHR (1 min) feature as the highest-ranking feature. This result is likely to be related to the notion that HR positively correlates with movement. However, in the same feature set, there were a few more features automatically selected in time-domain, frequency-domain and nonlinear indices including some of the features that best reflect the vagal tone.

While creating the FS3, HRV feature set, we first took the first 2 highest ranking features into consideration although they might be related to physical activity. Additionally, we focused on adding time domain features to our FS3 feature set which might give reliable assessment of vagal tone for the response windows shorter than 1 min. So, we added RMSSD and pNN50 time domain features into the FS3 feature set which were already highly ranked in the infogain algorithm. As a result, significant improvements were achieved by the Decision Tree algorithm, with 77.1 being the highest accuracy. We were capable of classifying the social initiation behaviors of ASC children during the Lands of Fog MR environment sessions and those occurring during the LEGO construction sessions based on psychophysiological data sources.

On the other hand, although the number of participants involved in this study was low, the number of initiations they have done in each experience have allowed for the creation of models for classification. This result seems promising, as we were unable to improve the model accuracy significantly, despite the number of initiations increased in each sub training set. However, further work is required to establish the viability of our model including the physiological measurements which are standardized according to the quantity of the movements.

Our quest is to understand the new potential that our MR system provides as an intervention tool, compared to current typical tools, e.g. possibly achieve a lower anxiety context. In this line of research, this study is providing a powerful model to differentiate how our users react differently in terms of psychophysiological arousal activity in both cases. The proposed computational model has already helped us see that certain HRV and EDA features which might be associated with social functioning, anxiety, or physical activity could have an effect on finding the differences between the social initiation taking place in the Lands of Fog condition and that taking place in the LEGO condition.This is of vital importance to make us move towards a better understanding on how a face-to-face, full-body interactive MR system, can become a good mediator of a socialization experience between ASC and non-ASC children and become a new tool for therapists and caregivers.

5.3 Body Cue Modeling of SIBs

[Parts based on the article *Sayis, B., Pares, N., Gunes, H. (2020, October). Bodily Expression of Social Initiation behaviors in ASC and non-ASC children: Mixed Reality vs. LEGO Game Play. In Companion Publication of the 2020 International Conference on Multimodal Interaction (pp. 140-149).*]

5.3.1 Introduction

In this study, our first goal was to show that an MR setting can be utilized to alter the non-verbal body behavior between ASC and non-ASC during social interaction as much as a traditional therapy setting (LEGO). A second goal was to show how these body cues differ between ASC and non-ASC children during social initiation in these two settings. We present our first analysis of the body cues generated under two conditions in a repeated-measures design. Body cue measurements were obtained through skeleton information and characterized in the form of spatio-temporal features from both subjects individually (e.g. distances between joints and velocities of joints), and interpersonally (e.g. proximity and visual focus of attention). We used machine learning techniques to analyze the visual data of eighteen trials of ASC and non-ASC dyads. Our experiments showed that: (i) there were differences between ASC and non-ASC bodily expressions, both at individual and interpersonal level, in LEGO and in Lands of Fog during social initiation; (ii) the number of features indicating differences between ASC and non-ASC in terms of nonverbal behavior during initiation were higher in Lands of Fog as compared to LEGO; and (iii) computational models evaluated with combination of these different features enabled the recognition of social initiation type (ASC or non-ASC) from body features in LEGO and in MR settings. We did not observe significant differences between the evaluated models in terms of performance for LEGO and MR environments. This might be interpreted as the Lands of Fog experience encouraging similar nonverbal behaviors in children, perhaps more similar than the LEGO environment, as the performance scores in the MR setting are lower as compared to the LEGO setting. These results demonstrate the potential benefits of full body interaction and MR settings for children with ASC.

5.3.2 Data

The data being used in this study was based on the reconstructed version of the “Event related responses dataframe” and included only the data from the last 18 trials and from both

ASC and non-ASC participants in social initiation moments with corresponding body cue measurements.

5.3.3 Experiments and Results

5.3.3.1 Statistical Hypothesis Testing

In this section, we present the details of the statistical hypothesis testing to determine if there are differences in the values of each body cue features for social initiation generated by ASC and non-ASC participants in both conditions (Lands of Fog and LEGO). Previous work found that children with ASC exhibit a clear deficit in movement observable from birth [Teitelbaum et al., 1998] and evident throughout life [Cook et al., 2013, Fabbri-Destro et al., 2009, Provost et al., 2007, Rinehart et al., 2006, Torres et al., 2013, Whyatt and Craig, 2013, De Jaegher et al., 2013, Downey and Rapport, 2012]. Nonetheless, it has also been shown that children with ASC can identify emotions in body language just as accurately as their non-ASC peers [Peterson et al., 2015].

Social communication includes both verbal and nonverbal communication, therefore when fostering social initiation skills in children with ASC with ICT we believe we must strive to incorporate nonverbal communication as well as verbal. Nonverbal behavior expression in children with ASC might have a role in the complexity of the social context. In MR environments, communication is not limited to verbal interaction and incorporates embodied interaction which goes hand in hand with social perception and social understanding. However, no research has yet assessed the impact of an MR full-body interactive system based on individual and interpersonal body cues in fostering social initiation in children with ASC, and compare it with a traditional intervention approach. In this regard, our hypotheses are formulated as follows:

H1: *There are statistically significant differences between ASC and non-ASC nonverbal (body) behavior at individual and interpersonal level in LEGO during social initiation*

H2: *There are statistically significant differences between ASC and non-ASC nonverbal (body) behavior at individual and interpersonal level in the MR setting during social initiation*

To validate these hypotheses, the Mann-Whitney U test was run as the difference in the values of each body cue features (averaged over all the initiations taking place during each play session for each participant) for the ASC and non-ASC. We present and compare the results of the Mann-Whitney U tests for LEGO and the Lands of Fog in terms of individual and interpersonal features. All the significance values reported in following result sections were at the significance level $p < .05$. As a next step, we develop classification models as a way to study the joint impact of aforementioned nonverbal body cues.

5.3.3.1.1 Results for H1 Based on the Mann-Whitney U tests with the target variable ASC/non-ASC, we found that several individual and interpersonal feature types were significantly different between ASC and non-ASC during social initiation in LEGO (see Table 5.3). Among the six angles obtained from triangles formed by: shoulder, hand and

Table 5.3: Types of features which differ between ASC & non-ASC during initiation in LEGO.

Feature Type	Feature Category
foot to nose	individual
angle right shoulder elbow hand	individual
upper joints velocity	individual
log-cov. of the joints distances (temporal)	interpersonal-prox

elbow; hip, shoulder and knee; hip, knee and foot, all considering left and right sides for individuals; an angle obtained from shoulder, hand and elbow was significantly different between participants during social initiation. Moreover, the skeleton joint distance between foot to nose was significantly different between participants during social initiation.The spatio-temporal feature type, the velocities and energy of the upper joints of the skeleton were also significantly different between social initiations of the participants. In terms of spatiotemporal proximity features 1 out of 120 features was significantly different between ASC and non-ASC social initiation during LEGO. We have not observed any visual focus of attention related features being significantly different between participants.

5.3.3.1.2 Results for H2 Based on the Mann-Whitney U tests with the target variable ASC/non-ASC, we found that several individual and interpersonal feature types were significantly different between ASC and non-ASC during social initiation in the Lands of Fog (see Table 5.4). Similarly to the LEGO condition we found that an angle obtained from shoulder, hand and elbow was significantly different between participants during social initiation. Additionally, the angle obtained from hip, knee and foot was significantly different between participants during social initiation. Again similar to the LEGO condition, foot to nose was significantly different between participants during social initiation. However, this type of individual difference also included nose to hip and distance between hand features in the Lands of Fog. In terms of spatiotemporal proximity features, 69 out of 120 features show significant differences between participants. We also observed differences between torso distance-based features. The minimum euclidean distance among all joints from individual one to the torso of individual two, and vice-versa, were significantly different between participants during social initiation. The obtained energy over the euclidean distances from all joints of skeleton one to the torso of skeleton two, and vice-versa were, also significantly different between participants during social initiation. Moreover, we have seen significant differences in head pose features expressed as yaw and roll between participants during social initiation. We found that among 102 individual features, four features for LEGO and five features for the Lands of Fog showed significant differences between participants during social initiation. On the other hand, we observed that among 251 (vfoa + prox) interpersonal features, 1 feature for LEGO and 81 features for the Lands of Fog showed significant differences between participants during social initiation. These results show that most of the features showing the nonverbal behavior differences between ASC and non-ASC were individual features in LEGO, while in the Lands of Fog these were interpersonal features. As a next step, we develop classification models as a way to study the joint impact of aforementioned nonverbal body features.

Table 5.4: Types of features which differ between ASC & non-ASC during initiation in MR

Feature Type	Feature Category
hand to hand	individual
angle right shoulder elbow hand	individual
nose to hip	individual
angle hip knee foot	individual
foot to nose	individual
vfoa roll	interpersonal-vfoa
vfoa yaw	interpersonal-vfoa
energy joints ind. to ind. torso	interpersonal-prox
ind. to ind. torso	interpersonal-prox
torso to torso	interpersonal-prox
log-cov. of the joints distances	interpersonal-prox
log-cov. of the joints distances (temporal)	interpersonal-prox

5.3.3.2 Automatic Classification

Automatic classification allows us to study the differences between ASC and non-ASC nonverbal behavior in detail for both conditions (Lands of Fog and LEGO). The problem is formulated as a 2-class classification task for both conditions. The initiations from ASC and non-ASC in Lands of Fog were named as ASCMR(n=117) and non-ASCMR(n=113). The initiations from ASC and non-ASC in LEGO were named as ASCLEGO (n= 121) and non-ASCLEGO (n=140).

We used the same features that were derived from the skeletons based on individual and interpersonal body movements (proxy & vfoa based social features). As a first step, all features were considered as a single feature set, as a combination of all the individual and interpersonal features, for each of the 2 scenarios (LEGO and Lands of Fog).

We used Weka Data Mining Software for automatic classification [Hall et al., 2009]. We used the accuracy, precision, recall and F1-score metric to evaluate the models. We compared the findings for classifiers widely used in automated human motion analysis, namely Random Forest (RF) and Support Vector Machine (SVM) [Gunes and Piccardi, 2008, Bernhardt and Robinson, 2007, Ajili et al., 2018]. The hyperparameters of each model are tuned based on two grids of parameters. The grid of parameters used for SVM model refers to; 1) kernel (radial basis function -rbf, the polynomial kernel), 2) penalty parameter C of the error term (5 values ranging from 0.1 to 1000), and 3) gamma Kernel coefficient for rbf (4 values ranging from 0.001 to 1). The best set of parameters leading to the best performance of the SVM model was obtained across a 10-fold cross validation. Moreover, the number of trees for our RF model was set to 500, performing 10-Fold CV with each combination of values. Table 5.5 and Table 5.6 depict the overall scores obtained for the feature set which consists of all the individual and interpersonal features, for each participant in a subject-independent manner for LEGO and Lands of Fog, respectively.

Table 5.5: Classification Performance of Different Algorithms on LEGO data

Method	Class	Precision	Recall	F-Measure	CV Acc
ZeroR	ASC	-	-	-	53.64%
ZeroR	non-ASC	53.60%	100.00%	69.80%	
RF	ASC	64.30%	52.10%	57.50%	64.37%
RF	non-ASC	64.40%	75.00%	69.30%	
SMO	ASC	**64.90%**	**59.50%**	**62.10%**	**66.28%**
SMO	non-ASC	**67.30%**	**72.10%**	**69.70%**	

Table 5.6: Classification Performance of Different Algorithms on MR data

Method	Class	Precision	Recall	F-Measure	CV Acc
ZeroR	ASC	50.90%	100.00%	67.40%	50.87%
ZeroR	non-ASC	-	-	-	
RF	ASC	60.20%	60.70%	60.40%	59.57%
RF	non-ASC	58.90%	58.40%	58.70%	
SMO	ASC	**60.90%**	**59.80%**	**60.30%**	**60.00%**
SMO	non-ASC	**59.10%**	**60.20%**	**59.60%**	

With automatic classification we were able to show that social initiation (ASC or non-ASC) recognition scores based on body features, as a combination of individual and interpersonal feature sets, for both LEGO and MR conditions are above chance level. Model scores were significantly different from the baseline performance (ZeroR). Moreover, we observed that classification scores were not statistically different between LEGO and Lands of Fog.

5.3.3.3 Video Analysis of the nonverbal behaviors between ASC and non-ASC

Following the analysis from the previous 2 sections, we also undertook a detailed visual analysis of social initiations of children with ASC and non-ASC in order to better understand the context around social initiation behaviors. The head and hands are seen responsible for the most movement and have received the highest attention in developing adequate coding strategies [Harrigan, 2013]. Therefore, we decided to focus on observing the movement of head and hands in LEGO which reflects the features that are significantly different between ASC and non-ASC in LEGO. MR environment involves a large-scale setting which lets players engage in collaborative activities by moving closer to their partner. Moreover, in body movement coding research the primary proxemic variable has been the distance between interactants, and it has garnered the most investigative attention [Harrigan, 2013]. Therefore, we decided to focus on observing the distance between participants in each social initiation moment from ASC and non-ASC in the MR environment which might cover most of the proximity based features discovered as significantly different in the previous sections. From 18 trials we randomly selected 4 trials (appx corresponding to %25 of all the trials) and from these 4 trials, a trained video coder, watched each social initiation video clip from ASC and non-ASC for LEGO and Lands of Fog.

5.3.3.3.1 LEGO As the upper joints velocity was significantly higher in non-ASC initiation moments compared to ASC initiation moments (Figure 5.3), we focused on hand and head movements. In the first 5 min of the sessions, searching the LEGO pieces from boxes was common for both participants which required a lot of hand movements and reaching out gestures to all boxes (Figure 5.4). However, there are less pieces on the LEGO table in the non-ASC initiation moments compared to the ASC initiation moments. This might signal that during non-ASC initiation moments, the participants are still in the searching process (Figure 5.4) and as a result they might have had more body movements. Moreover, there were more initiations (although it is not significantly different) in the first 5 mins of the session by the non-ASC children compared to ASC children (Figure 5.3).

5.3.3.3.2 MR As the torso to torso distance was significantly higher in non-ASC initiation moments compared to ASC initiation moments (see Figure 5.5), we focused on observing the torso to torso distances between ASC and non-ASC in those clips. We observed that during the initiation moments, when exploration is taking place in the playing plane, ASC children had a tendency to follow the non-ASC children more often and they were keeping a close proximity (see Figure 5.6). As it can be seen in the Figure 5.6 top image, the ASC (red net) kid initiates while he is following his non-ASC peer (blue net) during exploration. On the bottom image it can be observed that non-ASC discovers the virtual door under the virtual fog and shares this experience by creating a social initiation (pointing out to the door, calling the ASC kid) with the ASC kid (red net) who is at that stage outside of the playing field.

5.3.4 Discussion & Conclusion

5.3.4.1 Discussion

In our statistical hypothesis testing, we saw that the angle obtained from shoulder, hand and elbow, and the skeleton joint distance between foot to nose, were significantly different between ASC and non-ASC during social initiation in LEGO. In MR environment, the angle obtained from hip, knee and foot; and nose to hip and distance between hands, were significantly different between participants during social initiation. The aforementioned features are useful when discriminating between postures related to standing and sitting or torso inclination. We observed that more of these features play a role in differentiating between ASC and non-ASC during initiation in MR setting. This might be related to the diversity of full-body activities that the participants are motivated to undertake in MR setting.

A representative example of initiation in LEGO setting would involve changing the standing position which is likely to occur when picking up the LEGO piece dropped onto the floor. We also observed that the velocity and the energy of the upper joints were statistically different between ASC and non-ASC during social initiation in LEGO. This feature

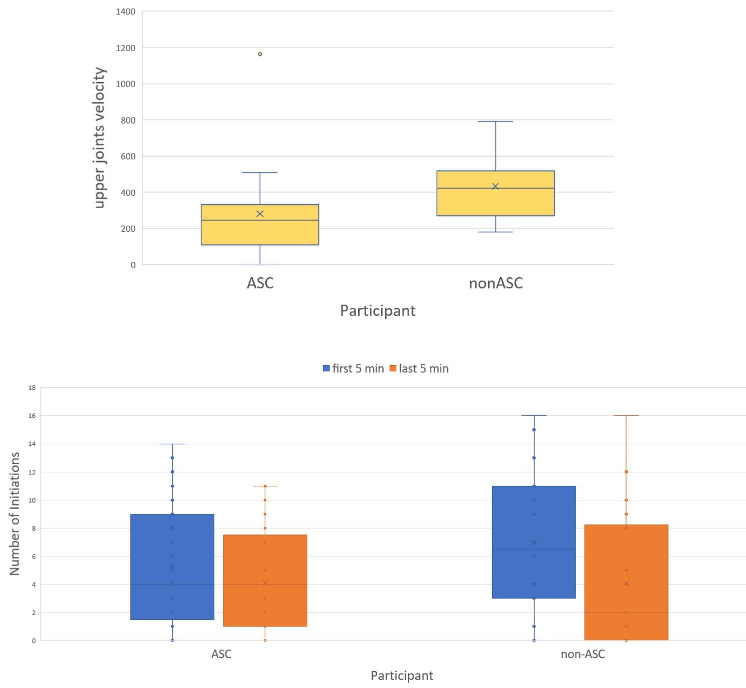


Figure 5.3: Top: upper joints velocity differences between ASC and non-ASC children in social initiation moments during LEGO (left). Bottom: number of initiations generated by ASC and non-ASC in the first and last 5 mins of the LEGO sessions (right).

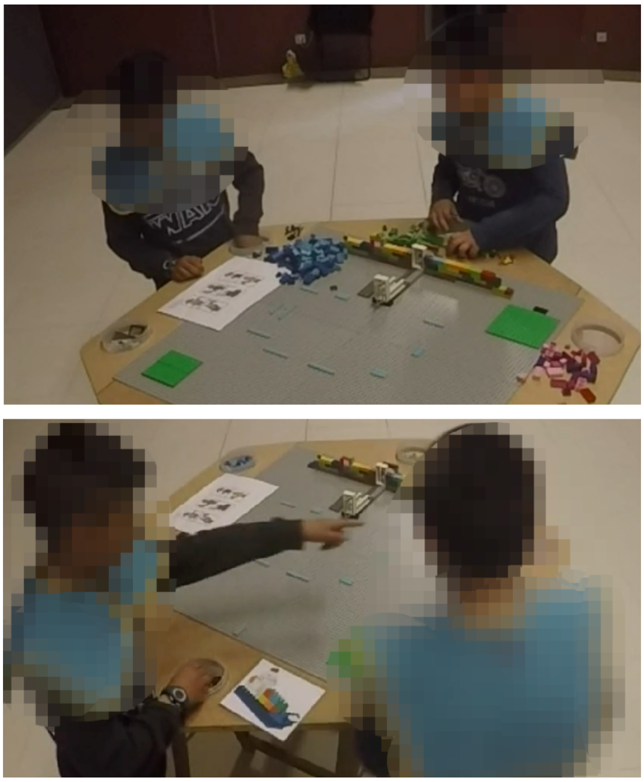


Figure 5.4: Top: non-ASC kid (on the left) and ASC kid (on the right) playing LEGO with social initiation by ASC kid. Bottom: non-ASC kid (on the left) and ASC kid (on the right) playing LEGO with social initiation by non-ASC kid.

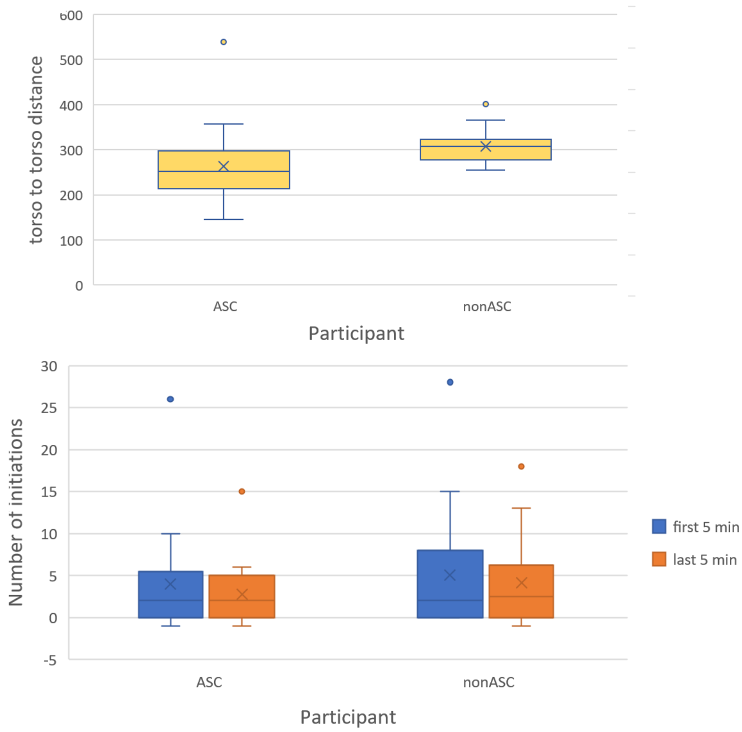


Figure 5.5: Top: torso to torso differences between ASC and non-ASC children in social initiation moments during the Lands of Fog experience (left). Bottom: number of initiations generated by ASC and non-ASC in the first and last 5 mins of the Lands of Fog sessions (right).

plays a role in the formation of key movement poses, e.g., in repeated movements, joints alternately accelerate and decelerate which leads to changes in the energy model representation. This information also helps the characterization of drastic changes in skeleton direction and velocity. In other words, in the LEGO condition, initiations by the non-ASC children, regarding the general upper joint velocity level is higher than the initiations from the ASC children; and this difference is statistically significant. On the other hand, there is no statistically significant difference between the initiations of ASC children with respect to non-ASC children in the Lands of Fog condition, regarding this general upper joint velocity. This seems to be an important finding, since our MR system with the Lands of Fog experience achieves similar nonverbal behaviors in both children, in contrast to the LEGO bricks (in terms of upper joint velocities). Although this needs further research in the future, it provides an encouraging advantage for our system compared to traditional interventions.

We also observed differences in torso distance-based features in MR. These features are important in identifying the most active person (i.e., the person entering into the other person's individual space). We also observed that majority of the features representing nonverbal behavior differences between ASC and non-ASC in LEGO setting were the individual features, while in the MR setting these were the interpersonal features. The large scale setting of the MR system might have led to higher numbers of locomotive actions, as players would display collaborative actions by moving within proximity to their partner, sometimes coming from across the play area. In such situations, players can accomplish

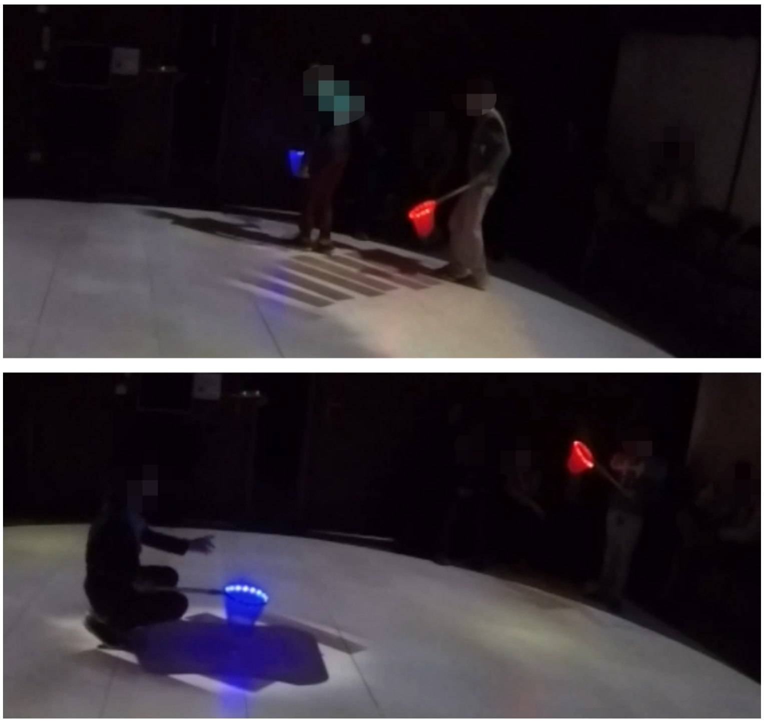


Figure 5.6: Top: non-ASC kid (with blue net) and ASC kid (with red net) playing in the MR environment and it is the moment of social initiation from ASC kid, Bottom: non-ASC kid (on the left) and ASC kid (on the right) playing in MR and it is the moment of social initiation from non-ASC kid.

joint actions without the need for verbal communication. As the LEGO setting provided a context where players were already within proximity to each other, sharing a focal point of joint attention, features related to differences in proximity and vfoa were not as commonly observed.

When analysing the computational models, the combination of all the aforementioned features allowed us to recognize the type of social initiation moment (ASC or non-ASC) from body cues in both environments, validating our hypotheses. Past research indicates that different body cues work together [Costa et al., 2001] and may contribute more information when treated together. In contrast to the analysis of some isolated features, here we did not observe significant differences between the evaluated models in terms of performance scores for initiations of the children in the LEGO condition with respect to the Lands of Fog condition. This might imply that the nonverbal behavior difference between ASC and non-ASC in Lands of Fog, based on all the features, is similar to the nonverbal behaviour difference between ASC and non-ASC in LEGO. Although this is not as good a result as was being hinted by some of the features, we can at least assume that our system is capable of encouraging nonverbal behaviours associated to SIBs in children with ASC with similar success as the traditional strategies, such as playing with LEGO bricks. Moreover, it must be noted that there is no human intervention in motivating the children in our system. This means that the interaction design and the mediation logic of the experience of our system is achieving a promising result without interfering in the children's activity. This is in contrast to the traditional strategies that are based on the abilities of the therapist to foster such behaviours in children and may potentially generate interference.

Moreover, despite the differences in both conditions are not statistically significant, the hinting that some of the features were providing of having more similar SIBs (of ASC vs non-ASC children) in the Lands of Fog condition than in the LEGO condition can still be seen since the difference, as per the computational model's performance, is smaller in the Lands of Fog condition than in the LEGO condition.

When we consider the manual video analysis for LEGO settings, the most common non-verbal behavior for both participants during initiation was related to searching LEGO pieces – i.e., hand movements and reaching out gestures. However, the number of LEGO pieces on the table for these initiation moments differed between ASC and non-ASC. This might signal that during non-ASC initiation moments, the participants are still in the process of searching the character figurines (Figure 5.4) and therefore they might have had more body movements enabling easier detection of nonverbal behavior differences. Moreover, searching activity for LEGO pieces might be related to imitation as both participants are doing similar activity. Past work on ASC demonstrated that motor imitation related impairments were common during social communication [Ingersoll, 2008, Rogers et al., 2003, Stone et al., 1997]. On the other hand, we observed that during the initiation moments, when exploration is taking place in the MR system play area, ASC children had a tendency to follow the non-ASC children and they would keep a close proximity. Data from several studies suggests that the ASC child usually has less chance to take a leading role in activities with mixed dyads of friendship. Therefore, settings that mediate communication between individuals should encourage balanced interactions to ensure that children with ASC have an equal opportunity in leading the interactions and the game [Bauminger et al., 2008]. As we were able to observe ASC initiation taking place at these close proximity moments, it can be said that Lands of Fog might be providing balanced opportunity for ASC children to take the lead. We also note that it was more likely to observe the non-ASC children exploring the MR system play area individually during non-ASC initiation. Evidence from previous comparison studies suggest that non-ASC children interact less with ASC children [El-Ghoroury and Romanczyk, 1999, Knott et al., 1995, Knott et al., 2007]. This might explain why non-ASC children are away from their ASC peers in non-ASC social initiation moments. In the MR condition, observing non-ASC children initiating is therefore promising in terms of fostering interaction between mixed dyads.

5.3.4.2 Conclusion

This study is a first step in analyzing the body behavior differences between ASC and non-ASC children during social initiation in a Mixed Reality (MR) environment and a typical LEGO-based intervention on social skills, where both environments mediate dyadic face-to-face play. To analyze the body cue differences between ASC and non-ASC, we extracted a set of individual and interpersonal spatio-temporal features derived from the skeletons of both subjects. Individual features included features such as euclidean distances between joints; angles formed between joints; and torso inclination. Interpersonal features included features which describe the relationship between two skeletons based on physical proximity, i.e. inter-body distance during the interaction and visual focus of attention. Extraction of these features allowed us to observe the specific individual and interpersonal features which differ between ASC and non-ASC children during social initiation in the MR and LEGO settings. In general, most of the features showing the nonverbal behavior differences between ASC and non-ASC were individual features in the LEGO setting and interpersonal features in the MR setting. Observing this difference for Lands of Fog is a promising result. Lands of Fog might have led to higher numbers of locomotive actions, as players would display collaborative actions by moving within proximity to their partner, sometimes coming from across the play area. In such situations, players can accomplish joint actions without the need for verbal communication. This shows that MR systems exploit full-body interaction, and incorporate navigation of physical and virtual space, take advantage of body gestures, allow relationships in space with the other user, and, in general, include nonverbal communication compared to other types of intervention.

We also investigated the joint effect of the various sets of features. As a result, we were able to automatically recognize the type of social initiation moments (ASC or non-ASC) from body cues in LEGO and in MR. Although there were differences between LEGO and Lands of Fog in some features in terms of finding the nonverbal behavior differences between ASC and non-ASC (as explained in previous paragraphs), when the complex relationship between all the features were investigated the computational models evaluated for social initiation classification did not show significant differences between the LEGO and MR settings. This is also a promising outcome indicating that an MR setting can be utilized to positively manipulate the nonverbal body behaviors between ASC and non-ASC children during social initiation as well as a traditional therapy setting (LEGO). However, further work is needed to discover hidden relationships between different body cues and their differences between ASC and non-ASC children.

Our manual video analysis confirmed that the body features utilised for automatic analysis were perceivable and meaningful to a human coder. However, these body features should be analyzed together with other data sources, such as system logs of game events, to have a better understanding of the differences between ASC and non-ASC social initiations. This will lead to a more insightful understanding of how a face-to-face, full-body interactive MR system can become a mediator to foster socialization between the ASC and the non-ASC children, and pave the way for designing novel tools that can be utilised by the ASC therapists and caregivers.

5.4 Multimodal Context-Driven Modeling of SIBs

[Parts based on the article *Sayis, B., Ramirez, R., Pares, N., (Submitted) n.d. Multimodal Context-Driven Evaluation of Social Interaction behaviors in children with Autism during Full-Body Interaction : A Comparison study between a Mixed Reality and a Nondigital Intervention.*]

5.4.1 Introduction

This study extends our previous research in which we showed that our Mixed Reality (MR) system fosters social interaction behaviors (SIBs) in children with Autism Spectrum

Condition (ASC). When comparing our system to a LEGO-based nondigital intervention we observed how it effectively mediates a face-to-face play session between an ASC and a non-ASC child providing new specific advantageous properties. A first goal of the present study is to obtain multimodal data models to extensively evaluate the overall effect of Lands of Fog condition, when compared to the LEGO condition. We have used context-driven approaches operationalised by (i) video-coding of SIBs, (ii) psychophysiological data, and (iii) system logs of user-system events. A second goal is to show how SIBs, taking place in these experiences, are influenced by the internal states of the users and the system. We measured SIBs by video-coding overt behaviors (Initiation, Response and Externalization) and with self-reports. We measured internal states using the wearable designed by our group to acquire: Electrocardiogram (ECG) and Electrodermal Activity (EDA) data. We used affective sliders and State Trait Anxiety Scale questionnaires as self-reports. In our repeated-measures design with two conditions, the MR environment and the traditional therapy LEGO, we used statistical inference analyses and machine learning techniques to analyze the multimodal data from 72 children (36 trials of dyads ASC with non-ASC child). Our results show that Lands of Fog condition has a positive effect on SIBs when compared to the LEGO condition, with an added advantage of being more flexible.

5.4.2 Data

The data being used in this study was based on the reconstructed version of the "Frequency of responses dataframe" and included the data from all the trials and only from the ASC participants with corresponding overt behavoir, systemlog and psychophysiological measurement data.

5.4.3 Experiments & Results

This study has analyzed how a full-body interactive ICT system, specifically an MR system, can be used to foster the SIBs in at least a similar degree as can be achieved with a traditional therapy setting (e.g. with LEGO blocks). Within this big picture, the study also analyzed how SIBs taking place in these experiences are related to the internal states of the participants as well as the events generated and sensed by the system. More in detail, the hypotheses of the work were:

Hypothesis 1: *Lands of Fog generates a similar frequency of SIBs (operationalised by the video-coding of their overt behaviors and questionnaires) in children with ASC as LEGO.*

Hypothesis 2: *The level of internal state activity during Lands of Fog sessions (operationalised as the psychophysiological measurements and questionnaires) will be related to the frequency of SIBs shown by the ASC children in at least the same level as they are related for the LEGO sessions.*

Hypothesis 3: *The count of Game Activity during the MR sessions (operationalised as the amount of events triggered by the system) will show a significant relation to the amount of SIBs shown by the ASC children.*

5.4.3.1 Statistical Inference Analyses

Our analyses took place in two phases. We first used paired sample T tests to evaluate the efficacy of Lands of Fog in affecting the SIBs as much as a traditional therapy setting (LEGO). We then analyzed the link between SIBs taking place in these experiences with internal state activity and game activity. We used correlation analysis, and multilevel linear regression analyses to evaluate their reciprocal influence. In our statistical analyses, we used SPSS 11 software [Spss, 2011].

5.4.3.1.1 Paired Sample T Tests

When we conducted paired sample T tests with the target variable MR/LEGO, results showed that there were no statistically significant differences in the operationalized measurements between the MR and LEGO conditions (Table 5.7).

5.4.3.1.2 Correlation Analysis

To better explore the possible link between SIBs with the internal state activity and game activity, we analyzed in both setups the correlations between SIBs and internal state activity; and the correlations between SIBs and game activity.

In LEGO, we found that positive correlations emerged between the internal state measurement RMSSD and Initiation, Response and Externalization. Moreover, a negative correlation emerged between the internal state measurement arousal_level and Externalization. Correlation analyses were not applicable between LEGO game activity measurements and other measurements as we have not logged the game events in LEGO (Figure 5.7).

In MR, there were no correlations between SIBs and internal state measurements. However, there was a positive correlation between social_status and number of collected fireflies during the MR experience (hunted_fireflies). Moreover, a positive correlation emerged between internal state measurement HRV (RMSSD) and game activity measurement distance between participants (distance_btw_participants) and a negative correlation between RMSSD and number of manipulated props (manipulated_props). Additionally, a positive correlation emerged between valence_level and distance_btw_participants. On the other hand, negative correlations emerged between valence_level and game activity measurements such as number of pointing at action of the virtual character towards other virtual objects (point_at_action) and number of virtual character merges between participants (merge_char) (Figure 5.7). As there were no correlations between SIBs and internal state measurements, in order to investigate any indirect relationship, we wanted to further investigate the relationship between hunted_fireflies and other game activity measurements which are associated with the internal state measurements. We found that there was a positive correlation (0.426*) between hunted_fireflies and distance_btw_participants (which was positively correlated with RMSSD).

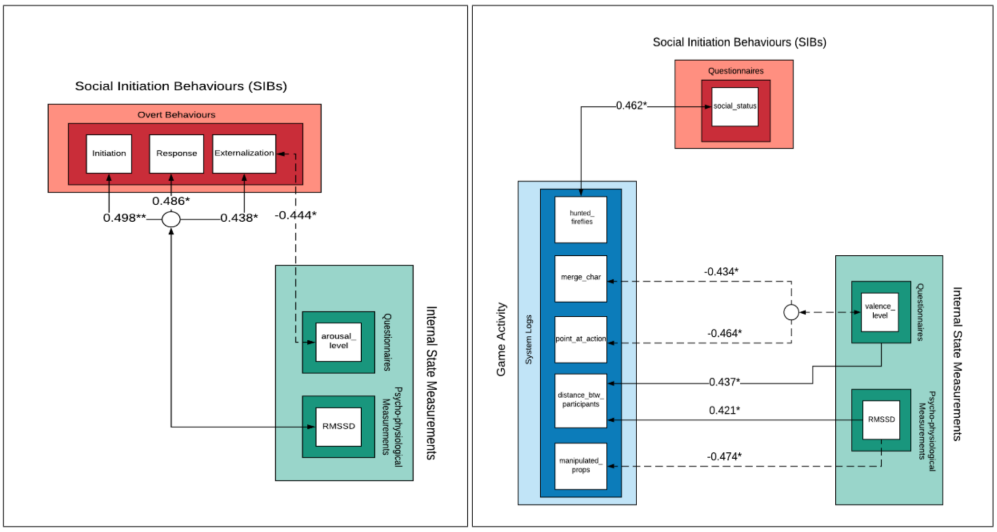


Figure 5.7: Left: Results for the correlation analyses in LEGO. Right: Results for the correlation analyses in MR. **Correlation is significant at the 0.01 level (2-tailed). *Correlation is significant at the 0.05 level (2-tailed).

5.4.3.1.3 Hierarchical Multiple Regression Analyses (HMRA)

In the analyses done for LEGO, SIBs from the video-coding were significantly predicted by some internal state measurements (Table 5.8, LEGO). In particular, the RMSSD was included in all the significant models with the only exception of the Externalization. Instead in the prediction of Externalization, CDAAmpsum and arousal_level were effective.

In the analyses done for MR, Initiation and Externalization were significantly predicted by some internal state measurements, game activity and SIBs from self report measurements (5.8 X, MR. In particular, desire_to_know_more and hunted_fireflies were included in all the significant prediction models.

Table 5.7: Paired Sample T tests with the target variable LEGO/MR

Modality	Item	LEGO	MR	Significance
Video Coding of Overt behaviors	Initiations	M = 9.39 (SD = 6.78)	M = 7.97 (SD = 7.32)	0.225
	Responses	M = 7.06 (SD = 5.06)	M = 5.48 (SD = 4.93)	0.084
	Externalizations	M = 5.61 (SD = 5.38)	M = 6.35 (SD = 6.01)	0.475
Psychophysiological Measurements	HRV (RMSSD)	M = -9.23 (SD = 23.82)	M = -10.81 (SD = 21.5)	0.705
	EDA (CDAAmpsum)	M = 70.29 (SD = 92.38)	M = 47.59 (SD = 65.97)	0.12
Questionnaires	social_status	M = 0.17 (SD = 0.33)	M = 0.27 (SD = 0.39)	0.32
	desire_to_know_more	M = 0.01 (SD = 0.1)	M = 0.02 (SD = 0.08)	0.609
	arousal_level	M = 0.05 (SD = 0.32)	M = -0.04 (SD = 0.36)	0.256
	valence_level	M = -0.01 (SD = 0.25)	M = -0.06 (SD = 0.22)	0.298
	STAIC_state	M = 39.34 (SD = 1.45)	M = 39.63 (SD = 2.25)	0.5

Table 5.8: Hierarchical Multiple Regression Analyses

LEGO

Predicted Variable	Models	Predictor Variable	Significance
Initiation	Model 1	RMSSD	0.006
	Model 2	RMSSD, social_status	0.003
Response	Model 1	RMSSD	0.012
Externalization	Model 1	arousal_level, CDAAmpsum	0.005

MR

Predicted Variable	Models	Predictor Variable	Significance
Initiation	Model 1	CDAAmpsum, desire_to_know_more	0.032
	Model 2	CDAAmpsum, desire_to_know_more,creature_changed	0.023
	Model 3	CDAAmpsum, desire_to_know_more,creature_changed, hunted_insects	0.002
Response	Model 1	desire_to_know_more	0.064
Externalization	Model 1	valence_level, social_status, STAIC_state	0.054
	Model 2	valence_level, social_status, STAIC_state, desire_to_know_more, hunted_insects	0.023

5.4.3.2 Automatic Classification

In order to better understand the deeper details of the link between SIBs with internal state activity and game activity, we brought all the aforementioned features together and also decided to analyze nonlinear connections between these features with ML techniques.

From the SIBs obtained from overt behaviors, as a first step, we wanted to analyze Initiation in detail since it is considered as one of the most important elements of social interaction [Sigman et al., 1999]. After all, social initiation is the primordial social act to improve social skills in general and lead a more autonomous life. We contextualized the levels of social initiation, based on the mean values of each interaction in each condition. Each participant's score was compared (lower or higher than) to the average number of initiations taking place in the experimental and control condition, and each participant's score was placed in one of the two classes generated, namely below average initiations or above average initiations.

In this regard, the problem is formulated as a binary classification task for both conditions. The below and above average initiations in Lands of Fog were named as in_LoF_below (n: 15) and in_LoF_above (n: 14). The below and above average initiations in the LEGO experience were named as in_LEGO_below (n: 15) and in_LEGO_above (n: 14).

In order to combine the information from all the unimodal features, we generated a single fused vector integrating the data from all the modalities. The resulting feature sets for LEGO (64 features in total, 47 HRV, 12 EDA, 5 Questionnaires) and MR (74 features in total, 47 HRV, 12 EDA, 5 Questionnaires, 10 Game Activity) were quite large. Therefore, to reduce dimensionality of the data to improve the classification performance we used different feature selection techniques.

Using Weka Data Mining Software [Hall et al., 2009], we compared the results of Neural Networks (NN), a state-of-the-art nonlinear classifier, and Decision Trees, and interpretable classifier to better understand the modeling process. Neural Networks are seen as one of the most effective techniques to model complex nonlinear hypotheses even when the input feature space is large. They are not bounded by assumptions about the features distributions. However, Neural networks were recently viewed as 'black boxes' as they could not explain their predictions. On the other hand, decision trees are very intuitive data structures, and visualizing them can give valuable insights on their criteria for classifying new data.

5.4.3.2.1 Feature Selection

We used two types of feature selection strategies:

All Features (FS1): For LEGO (64 features, 47 HRV, 12 EDA, 5 Questionnaires) and for MR (74 features, 47 HRV, 12 EDA, 5 Questionnaires, 10 Game Activity) were used as a result of feature extraction.

WrapperSubsetEval (FS2): A wrapper approach (with a best-first search method) to feature subset selection was used in order to find a minimum set of attributes that are tailored to the particular domain which maximize the performance of the classifiers. The results of the feature selection strategies can be found in Table 5.9.

5.4.3.2.2 Algorithms

The study includes an overall comparison between the models obtained with two feature selection strategies (FS1,FS2) for both conditions. We used the accuracy, precision, recall and F1-score metric to evaluate the models. The hyperparameters of each model are tuned based on two grids of parameters. The grid of parameters used for the NN model refers to a learning rate ranging from 0.0025 to 0.1 with the step size 0.0025 and momentum ranging from 0.1 to 0.5 with the step size 0.1. The grid of parameters used for the Decision Tree model refers to the confidence factor ranging from 0.1 to 0.9 with the step size 0.1 and minimum number of objects ranging from 1 to 4 with the step size 1. The best set of parameters leading to the best performance of the models were obtained across a 10-fold cross validation. Table 5.10) depicts the overall scores obtained for the feature sets FS1 and FS2, for each participant in a subject-independent manner for LEGO and Lands of Fog, respectively. With automatic classification we were able to show that social initiation level recognition scores based on internal state activity and game activity, as a combination of feature sets, for both LEGO and MR conditions are above chance level (ZeroR CV accuracy: 51.72%). Model scores were significantly different from the baseline performance. Moreover, we observed that classification scores were not statistically different between LEGO and Lands of Fog.

Table 5.9: Feature Selection Strategies

Category	Type	Dataset
All (FS1)		
Time-domain, Frequency-Domain, Nonlinear Features Continuous Decomposition (CDA), Standard trough-to-peak (TTP) Analysis STAIC_state, arousal_level, valence_level, social_status, desire_to_know_more	HRV EDA Questionnaire	LEGO
Time-domain, Frequency-Domain, Nonlinear Features Continuous Decomposition (CDA), Standard trough-to-peak (TTP) Analysis STAIC_state, arousal_level, valence_level, social_status, desire_to_know_more manipulate_prop, merge_char, point_at_action, hunted_fireflies, distance_btw_participants, time_of_first_char, char_greeting, char_texture_change, char_being_idle, area_covered	HRV EDA Questionnaire Game Activity	MR
Wrapper Subset(FS2)		
pNN50(%),HF(ms2̂), HF(n.u.), HF(%), VLF(Hz), LF(ms2̂), Meanlinelength(beats)	HRV	LEGO
TTPnSCR, CDASCR DET, LF/HF manipulateprop, time_of_first_char desire_to_know_more	EDA HRV Game Activity Questionnaire	MR

Table 5.10: Classification Performance of Different Algorithms on LEGO and MR data

LEGO	FS1				FS2			
Method	Precision	Recall	F-Measure	CV Acc.	Precision	Recall	F-Measure	CV Acc.
Decision Trees	24.1%	24.1%	24.1%	24.1%	69.6%	69.0%	68.5%	**68.9%**
Neural Networks	66.0%	65.5%	65.4%	65.52%	89.8%	89.7%	89.6%	**89.6%**
MR	FS1				FS2			
Method	Precision	Recall	F-Measure	CV Acc.	Precision	Recall	F-Measure	CV Acc.
Decision Trees	62.3%	62.1%	61.5%	62.0%	84.3%	82.8%	82.6%	**82.7%**
Neural Networks	51.7%	51.7%	51.7%	51.7%	89.9%	89.7%	89.7%	**89.6%**

5.4.4 Discussion & Conclusion

5.4.4.1 Discussion

5.4.4.1.1 Hypothesis 1: Lands of Fog generates a similar frequency of SIBs in children with ASC as LEGO

The video-coding of overt behaviors revealed no significant differences in the generated number of SIBs between the MR setting and the LEGO control condition. This finding was also consistent with our findings from automatic classification when we wanted to analyze the complex dynamics between all the features which contribute to the generation of different levels of SIBs (initiation). We did not observe significant differences between the evaluated automatic classification models in terms of performance for LEGO vs. MR environment. These results imply that our system is capable of fostering SIBs in children with ASC with similar success as the traditional strategies, such as playing with LEGO bricks. This is important since there is no human intervention in motivating the children in our system, whereas traditional strategies depend on the abilities of the therapist to foster SIBs in children. This does not mean that we propose an unsupervised system that does without experts. On the contrary, our system can, on the one hand, free the experts from having to mediate the session and allows them to focus on observing the evolution of the child within the intervention program. On the other hand, our system fosters SIBs in ASC children in a homogeneous and unified manner for all children without incorporating human subjectivity. The presence of therapists has been found to interfere in traditional sessions since they represent an added human actor in the activity. Because of this, they cannot be absolutely sure whether the social event generated by the ASC children is an initiation that comes from their own inner drive and will, or whether it is merely a response to previous actions started by the therapists themselves. In an MR experience, however, the system defines a context that mediates the session between the children without having the human figure get in the way. The experts can define the parameters of the MR system to adapt the session to each child and can intervene in the session only if they deem necessary.

Analyzing the questionnaires, we saw that no significant difference existed between the two conditions regarding the reported level of knowledge of the play peer (social_status), and the desire to know more about the peer after playing (desire_to_know_more). In this regard, the promotion of collaboration in the MR environment is probably helping ASC children to see their peers as valuable play partners, in a similar way to how they see peers in the LEGO condition. Given the challenges in social integration and maintaining social relationships in children with ASC [Strain, 1983, Gupta et al., 2014], creating scenarios where users see each other as valuable partners could help in developing and maintaining friendship. Lands of Fog experience offers the possibility of having a broad diversity of scenarios and can therefore provide even more possibilities to create collaboration situations, compared to a static unchanging approach such as the LEGO bricks are.

5.4.4.1.2 Hypothesis 2: The level of internal state activity during MR sessions will be related to the frequency of SIBs shown by the ASC children in at least the same

level as they are related for the LEGO sessions.

Based on the Paired Sample T tests with the target variable MR/LEGO, we found that STAIC_state results were not significantly different between MR and LEGO. While we have seen above that we are fostering SIBs in the MR condition, here we see that the children's perceived state anxiety, assessed through the STAIC_state questionnaire, was similar to that of the LEGO condition. This is an important finding, since Lands of Fog is a new and unknown play context, while the LEGO condition is a very well known and familiar play context. The LEGO play context in general, tends to make the children either play solo building their own thing, or play together with the pieces from the start if the therapist medites the experience and motivates children to build something together. On the other hand, Lands of Fog provides different levels of relationship between users and the system. Lands of Fog allows the users to start by doing solo activities and slowly and seamlessly brings the two together to allow them to discover joint play situations. Moreover, Lands of Fog allows the children to step back from collaborative play and go back to solo playing at any one time, and then back to collaboration again. We believe that this design approach (named encouraged collaboration, as opposed to enforced collaboration which is most often used in ICT for ASC), which gently encourages collaboration, created a comfortable setting for socialization and hence lowered potential anxiety coming from a novel and unknown situation.

On the other hand, the results obtained for the reported arousal/valence levels of children after playing show no significant difference between the two conditions. There is considerable evidence that factors specific to anxiety include negative valence as well as positive arousal, as mapped on the circumplex model of affect [Feldman, 1995]. In this regard, not seeing any significant difference in reported arousal/valence measurements between the MR and LEGO condition might also be related to the results we found from the state anxiety scales. However, we must be careful with the fact that the reliability of self-reports from children with ASC might show fluctuations as they may have difficulties in interpreting their own experiences [Berthoz and Hill, 2005].

Psychophysiology, as an objective measure, might provide more representative information about the children's internal state. Hence, we monitored their physiology and our results from both HRV and EDA did not exhibit statistically significant differences between MR and LEGO conditions. In our case, we see that results from physiological analysis seem to support the self-report analysis described above. Firstly, it should be noted that HRV indicates the parasympathetic nervous system (PNS) index and the PNS activity through vagal tone is associated with better social functioning [Laborde et al., 2017]. Seeing both conditions induce a similar level of HRV might show that the choice of an MR full-body interactive system has an effect on socialization as much as a traditional setting. Besides, seeing positive correlation between the level of HRV and SIBs in LEGO might also validate the association between HRV and socialization mentioned before. However, we needed to understand the direction of this association to be aligned with the previous findings. In this regard, Hierarchical Multiple Regression Analyses (HMRA) showed that with the level of HRV (RMSSD), we were able to predict the frequency of SIBs (initiation and response) in LEGO. It can therefore be assumed that the higher HRV can be associated with higher frequency of SIBs in LEGO. These results were also consistent with our findings from automatic classification where we analyzed the complex dynamics between all the features (not just with HRV and few other features as in HMRA) which contribute to the generation of different levels of SIBs (initiation). Firstly, with automatic classification we were able to show that social initiation level recognition scores (based on internal state activity, game activity and questionnaires, as a combination of feature sets) for both LEGO and MR conditions were above chance level. Moreover, model scores were significantly different from baseline performance. Secondly, the discovered features with the wrapper subset evaluation method (FS2) just included HRV-related features in the classification model for the LEGO condition. However, on the contrary, in the MR environment we have not seen a similar relationship (neither in correlation analysis, or in HMRA analysis, or in Automatic Classification) between HRV and SIBs. This might be because of the sensitivity of the HRV measurements in scenarios demanding physical activity. It is important to bear in mind that when Autonomic Nervous System (ANS) activity is assessed during exercise through HRV, the analysis of time and frequency domain measures of HRV may not yield adequate information [Boettger et al., 2010]. In this regard, although we have designed the two conditions as similar as possible, physical activity might be higher in the MR condition and hence, still getting in the way of understanding how HRV-related activation is associated with the SIBs in the MR environment (because of the allowed range of full-body activity). Perhaps, this is why we do not see a correlation between HRV and SIBs, or any HRV-dominant features in predicting the levels of SIBs in the MR environment.

On the other hand, EDA is often considered one of the most useful indices of sympathetic nervous activity (SNS) [Boucsein, 2012] and might provide different insights compared to HRV [Posada-Quintero and Chon, 2020]. This finding was reflected in our results as well. We saw that an EDA feature (CDAAmpsum) was more effective than a HRV feature (RMSSD) in predicting the levels of SIBs (initiation) in the MR environment in HMRA based analyses. Moreover, EDA might be related to engagement in addition to state anxiety [Hernandez et al., 2014]. Thus, seeing similar levels of EDA for both conditions can be associated with a similar response of children to externally presented stimuli in both conditions. Moreover, our results show that desire_to_know_more is coupled with an EDA measurement in the HRMA and in the Automatic Classification based prediction models of SIBs (initiation) in our MR environment and not in LEGO. This suggests that our MR environment is probably creating an adequate engaging environment for getting to know the peer and hence foster generation of SIBs.

5.4.4.1.3 Hypothesis 3: The count of Game Activity during the MR sessions will show a significant relation to the amount of SIBs shown by the ASC children.

In MR, there were no direct correlations between SIBs and internal state measurements of the children. However, there was a positive correlation between social_status and number of collected fireflies during the MR experience (hunted_fireflies). It could be argued that there might be an indirect relationship between internal states and SIBs as we also found that there was a positive correlation between hunted_fireflies and distance_btw_participants (which is a feature related to some of the internal state measurements). A possible explanation for this three-way relation can be as follows.

First of all, hunting fireflies was a simple introductory game mechanic that children could easily understand and perform alone, as a first "low floor" [Papert, 1980] contact with the virtual environment. However, our interaction design also envisioned this as opportunities to get the two children interested in one another through "search & discovery". For example, if a child is catching many fireflies and the other cannot, then the child with less fireflies can have sufficient interest to ask the other how he has caught so many. Another example would be when catching fireflies actually leads to the climax with the caught fireflies magically fusing into a character. This can make the child that achieves this exclaim his surprise and attract the interest of the other, or want to share emotions and thoughts. Hence, despite this being designed as a solo activity, the use of this simple introductory mechanic seems to have been successful in making children engage with the peer, as results show an increase in children's social_status.

Secondly, the positive correlation found between the distance_btw_participants and hunted_fireflies might also signal that a larger distance between participants might lead players to search the playing area more extensively, and hence collect more fireflies and discover different surprising elements. Thirdly, we must be careful with the fact that non-ASC children might feel more comfortable in approaching others, compared to ASC children [Gessaroli et al., 2013]. Thus, the unnegotiated approach of the non-ASC child towards the ASC child might have created some level of discomfort in the ASC child. This can be interpreted from the following findings: (i) positive correlation between distance_btw_participants and RMSSD; and (ii) positive correlation between distance_btw_par ticipants and valence_level. Such a discomfort is not found in the LEGO environment because this setting provides a context in which the players were already within proximity of each other, sharing a focal point of joint attention on the table (despite them having to move around it). However, once the barrier in close proximity is achieved in the MR environment (when the approach is negotiated), there might be positive correlations between RMSSD and SIBs in MR as well (as long as the physical activity does not get in the way of the measures), as in LEGO. It is therefore likely that such connections exist between SIBs and internal state measurements of the children in MR.

Additionally, character merging (merge_char) might also show the presence of collaborative behaviors. A significant increase in the occurrence of these might indicate the game's potential to foster SIBs. However, the obtained negative correlations emerging between valence_level and the merge_char might be a sign of a weak implementation of this design aspect of the MR setting. One evidence of this is that we have observed in several video recordings that our system was activating the character merge action when the participants' creatures were close enough to each other, although the participants' intentions seemed to be directed towards other activities. It is possible that the low contingency of these merge events might have led to a negative experience which is represented in the valence_level. We must make sure in future designs that no actions are triggered by the system unless it is clearly activated by the users.

Similarly, we have observed negative correlations emerging between valence_level and the character "pointing at" gesture (point_at_action). Having the character pointing at something was designed to help the users discover the nearby elements which were interactive, hence it is activated when the character of a user approaches the virtual elements. At that moment, the user's creature points towards the element and makes an exclamatory remark. However, this action might be disruptive (probably reflected by low valence_level) for the child; e.g. imagine the child was trying to move his/her character towards a specific zone of the MR system play area and suddenly the character stops following and does something unexpected. This could be an annoying situation for the child leading to the low valence we have observed.

On the other hand, we have also observed how the collaborative act of manipulating virtual objects (props) through the characters of the children might have had a positive effect on them resulting in better social interaction. Aforementioned possible indirect relationships between SIBs (initiation) and internal state measurements could be observed from the prediction model of the initiation (model 3) derived from HMRA analysis. When the hunted_fireflies feature was added to the prediction model 2, which already included an EDA-based feature, we could see an improvement in the significance of the prediction. Moreover, in automatic classification models, the possible indirect relationships between SIBs (initiation) and internal state measurements could also be discovered but in a more complex manner (in terms of predicting the level of initiation). Past research has indicated that different physiological measurements work together [D'Mello and Kory, 2012] and may contribute more information when treated together. Besides, when the feature selection techniques were considered (e.g FS2), the classification scores improved. Although it was not possible to interpret this complex relationship with the NN models, we were able to see that Decision Trees provided interpretable results with the feature manipulate_prop and EDA feature (TTPnSCR). TTP.nSCR (Number of significant skin conductance response (SCRs) within the response window) is also another EDA feature which represents the phasic activity (http://www.ledalab.de/documentation.htm.) and can be related to the external stimuli from Lands of Fog experience and engagement [Hernandez et al., 2014]. Manipulate_prop is the measure of the collaboration act in the MR environment as the virtual elements may only be manipulated when both partners work together. The decision tree structure that we have obtained is as follows: We observe the number of social initiations above average in the ASC child only when the TTP.nSCR is higher than a certain threshold (namely, 138) and the feature Manipulate_prop is greater than three (i.e. 3 manipulated objects). In any other case, we observe below average social initiations. In this regard, the mechanism of manipulating props seems to be a significantly adequately designed game mechanics to bring the children together and get them to collaborate and become interested in the other. And that the more props they manipulate, the better their relationship becomes. With an apparent threshold of 3 as a minimum to really get the initiations up above average. But also that to achieve this we need a good previous engagement of the children in the game. This could be related to many factors that need to be reassessed in the future. But for now we know that if the children are engaged (seen by TTP.nSCR>138) then they have more possibilities of manipulating sufficient props to get involved with the peer and therefore get the initiations above average. However, one might argue that the explanation previously given around the negative correlation found between point_at_action and valence_level might be linked to the findings from the decision tree as having the character pointing at something could also lead discover potential objects to manipulate, since it is activated when both users approach the virtual elements together. Therefore, it could be argued that possible failure in the manipulating a prop, when the ASC child attempts this in a solo action, might have caused some frustration, which might explain also part of the low level of valence. However, the negative valence correlated with pointing of character might be more probably associated with the fact that the character stops following the child and does something unexpected as explained before. This perhaps would tell us that in the future, the character should only point at props when the child (and the character) are in an idle situation rather than the character doing this action all the time. Then the character would be introducing a useful queue. The other feature which was effective in automatic classification models was time_of_first_char. This feature is related to the number of hunted fireflies in the first minutes of the game. After hunting a certain amount of fireflies, it was decided that fireflies would transform into a creature that would become a virtual partner for the user. As we discussed before, hunting fireflies was a simple introductory game mechanic which children could do alone and easily understand as a first contact with the system. Hence, despite the time it takes to collect the first necessary amount of fireflies, being designed as a solo activity, the use of this mechanic, when it is considered together with other features, seems to have been successful in making children engage with the peer in different social initiation levels, as results show above baseline prediction scores in the prediction of the level of social initiations.

5.4.4.2 Conclusion

This study provides new insights into the results from previous studies that indicated the potential of our mixed reality (MR) system in fostering social interaction behaviors (SIB) in children with Autism Spectrum Condition (ASC). Moreover, the present study has been one of the first attempts to thoroughly examine how a full-body interactive MR system can be used to foster the SIBs in at least a similar degree as can be achieved with a traditional therapy setting (e.g. with LEGO bricks). Within this big picture, the study also analyzed how SIBs taking place in these experiences are related to the internal states of the participants as well as the events generated and sensed by the system.

The results of this research support the idea that our system is capable of fostering SIBs in children with ASC with similar success as the LEGO setting, with an added advantage of being more flexible. This finding reported here shed new light on developing a tool that is mediating, guiding, and supporting the progress of the children in terms of practicing SIBs; and providing structure and assistance to therapists. Moreover, a further longitudinal study could assess the long-term effects of this tool, especially when it is used in place of a passive social skills training program that needs to be mediated by a therapist.

The results of this study also indicate that SIBs are related to some of the internal state measurements suggesting similar levels of anxiety levels, similar levels of social functioning through HRV and similar levels of engagement through EDA-based measurements in both conditions. Nevertheless, physiological measures are susceptible to physical activity (especially in our case HRV). Indeed, special care must be taken although we have designed the two conditions as similar as possible. Perhaps, for future work, physiological measurements could be standardized with the quantity of body motion information from the participants, which could be tracked with special algorithms through camera systems. Additionally, as a next step physiological measurements can be coupled with computer vision based internal state evaluation (e.g face and body gesture recognition) for more accurate evaluation. Implementation of such multimodal objective evaluation tools can pave the way for carefully designing and assessing the novel tools that can be utilised by the ASC therapists and caregivers.

Finally, the results of this investigation also show that SIBs are related to some of the game activity measurements. The results have provided very useful insight on which game mechanics were useful and which were not. With this information we can improve the MR experience for the children, both by improving the current game mechanics, as well as designing new ones that can appeal to a broader range of children in the autism spectrum. We have also seen that basing our design on the encouraged collaboration approach is helpful in fostering SIBs. One of the important findings is that we can keep up engagement in children by providing sufficient interaction potential, using game elements and mechanics such as hunting fireflies and manipulating props collaboratively. Also that we allow the children to step back from collaborative play and go back to solo playing at any one time.

We must continue the analysis of these full-body interaction MR systems to better help ASC children in their social skills, as well as helping non-ASC children better integrate ASC children into society.

Chapter 6

EXPLORING SOUND IN FULL-BODY INTERACTION

6.1 Introduction to this chapter

This chapter gives information on the methodologies used to explore new interaction design techniques through sound in order to contribute to the understanding of how full-body interaction environments can be improved. It presents an exploratory study and describes the design process, the development and the evaluation of a full-body interactive soundspace (GenPlay Soundscape). With this additional research, we demonstrate that full-body interaction is a suitable approximation for developing systems to engage users and motivate them to create SIBs with the help of sound.

This study is part of a larger project called Generative Playscape (GenPlay) that aims to supply novel resources to the teachers of integrated classroom settings which would serve as a support for increasing inclusion during playtime activities. Our work associated to this project was chosen among the top 12 entries in the Student Design Competition of the 2018 CHI Conference on Human Factors in Computing Systems (CHI 2018)

6.2 Background

Children's language/speech skills have been found to benefit from music interventions [Overy, 2003, Tallal and Gaab, 2006]. Moreover, in terms of increasing social responsive behaviors, it has been shown that music interventions have significant improvements over non-music interventions [Finnigan and Starr, 2010]. Additionally, there is a growing body of literature that recognises the potential of music therapy for individuals with ASC [James et al., 2015]. Children with ASC considers participating in musical interventions as a fun activity. Musical activities may improve their social interaction capabilities, thereby promoting communication and social skills development. While individuals with ASC show difficulties in social communication, it has been shown that their way of perceiving the music was similar to non-ASC individuals [Heaton et al., 1999] - children with nonverbal ASC, for example, are still able to sing [Wan et al., 2010]. Approaches including musical improvisation, where the interactive use of music does not represent a fixed or specific context, appear to draw interest and engage children with ASC, and orient them towards social interaction behaviors [Overy and Molnar-Szakacs, 2009].

Although children with ASC find the motivational drive to participate in musical activities, learning how to play a traditional musical instrument might be challenging. The use of full-body interaction environments can address these obstacles and provide a viable solution to create interactive musical activities directed towards children with ASC. Full-body interactive environments can help overcome apparent barriers in motor activity and allow generation of rich and natural musical expressions [Sigrist et al., 2015, Samuels, 2014], Several attempts have been made to demonstrate how music bonds to movement [Srinivasan et al., 2016]. Leman discusses in detail the role of synchronization, attuning, and empathy and their coexisting nature in the context of music-induced corporeal expression [Leman, 2008]. Forti et al presents an intervention for children with ASC based on imitation of corporal movements supported by a musical feedback [Forti et al., 2020]. Bendablesound is an another project which promotes motor skills and coordination through an elastic multimodal interface using sound [Ortega et al., 2015]. Recent work by Ragone et al. with the project name OSMoSIS supports children with ASC in generating sounds through free body-movements. [Ragone, 2020].

These ideas were taken into account when designing GenPlay Soundscape which specify a discrete and structured soundspace, and which could be experienced by children with ASC to promote exploration and foster social interaction between peers.

6.3 Research Framework

6.3.1 GenPlay Soundscape

The methods described in this thesis (GenPlay Soundscape) have been designed to test the efficacy of the sound interactions designed and developed to be part of the GenPlay system.

6.3.1.1 Setup

GenPlay is a mobile MR system (2 x 3 x 1 meters) allowing full-body interaction within a circular floor projected Virtual Environment with a soundscape where both children with ASC and children with non-ASC can explore, discover and engage with sounds and visuals (Figure 6.1). It is inspired by the interaction paradigms of Lands of Fog (6 x 6 x 4 meters) but tailored to a classroom setting. GenPlay's configuration is based on a Kinect (second generation) sensor which allows detection of full-body movements in 3D space with frontal motion capture approach in contrast to Lands of Fog's ceiling based approach which requires a higher space for installation. It allows mapping skeleton and small gestures with high accuracy. GenPlay's configuration also includes a loudspeaker set for immersive sound. GenPlay's both visual and sound interaction mixed reality environment was developed with Unity.

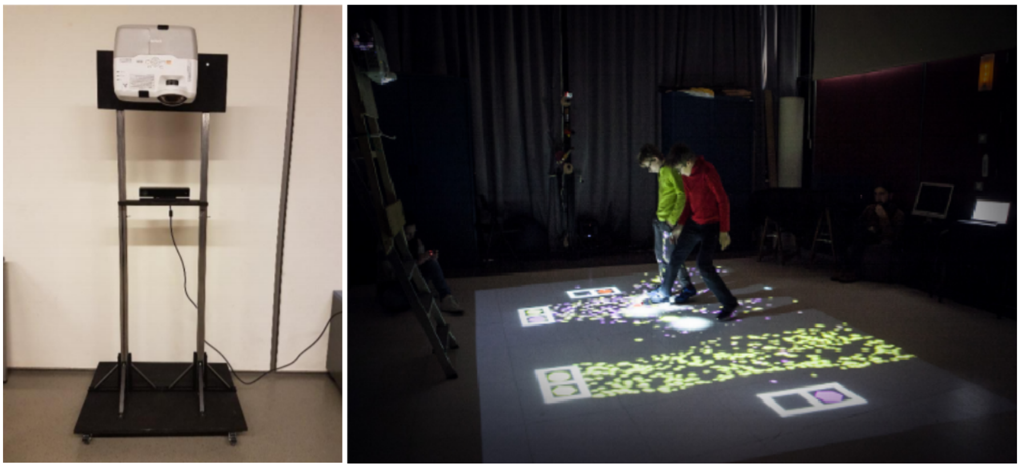


Figure 6.1: A photo of the GenPlay physiacal module (left) / Two children playing an initial system iteration (right).

6.3.1.2 Content & Interaction Design

GenPlay's audiovisual content and interaction dynamics have been designed and developed in two parallel phases namely GenPlay's Isolated Generative Sound Interactions and GenPlay's Isolated Generative Visual Interactions. Later, these two phases were integrated into one. As we have tested the effectiveness of GenPlay's sound elements compared to GenPlay's visual elements, we also briefly mention these visual elements in this thesis.

6.3.1.2.1 Sound Elements

The architecture of GenPlay's Soundscape follows the same processes commonly undertaken by a music therapy session: a first step of increasing awareness of human and musical relationship involved in the experience and a second phase of directing the participant to improvise and find a suitable and appropriate means of self-expression at his/her own pace [Gold et al., 2006].

In the first phase, it is planed to familiarize the children with the environment by using spatial positioning. The children can locate themselves in the floor's circular play area which is segmented into 24 parts, and each part is associated with a different note. Changing locations triggers notes to be played one after another and this composition fades out when the child stands still. In order to make the combination of sounds pleasant, sounds were decided to be categorized within the same scale which is C Major. It was also decided that each user can play the sounds of one of two musical instruments. For example, the sound of a piano playing three notes in a sequential manner (MI-SOL-TI) would be triggered as participant one moves from one position to another. On the other hand, the sound of a guitar playing two sequential notes separated by an octave (DO4-DO5..) would be triggered as participant two moves from one position to another. The reason we chose the generation of such collection of sounds as opposed to individual notes were based on the increased activity level we observed from our initial tests with testing participants. We also considered mapping tempo in user's interaction. We were inspired by Castellano et al.'s proposed strategy where the participant's full-body emotional behavior in the space is directly mapped to the emotion in the resulting performance [Castellano et al., 2007]. In this regard, the time frame between each combination's note was chosen to be shorter (faster tempo) or longer (slower tempo) corresponding to how fast or slow the participant is moving (Quantity of Motion). The notes are arranged so that if the child walks in a circular trajectory, the first note of the combination would eventually sound in the scale to which it belongs to, and the progression would be in arpeggiated chords if they walk towards the center of the playing area. This structure can be seen in Figure 6.2 (It should be noted that the diagram seen in this figure is not providing a representative structure of the GenPlay's visual environment). Moreover, we were also inspired by the findings of Jaschke's (2014) [Jaschke, 2014] study which shows that structural associations between sound environments and mechanical structure of music can help reduce hyper-sensibility among children with ASC. This was also supported by the Auditory Perception Training with Neurological Music Therapy technique which focuses on auditory perception and sensory integration, in this case kinesthetics[Thaut and Hoemberg, 2014]. In this regard, the displacement around the floor area was also assumed to be a relevant interaction design item which could help the children to identify and discriminate between different notes.

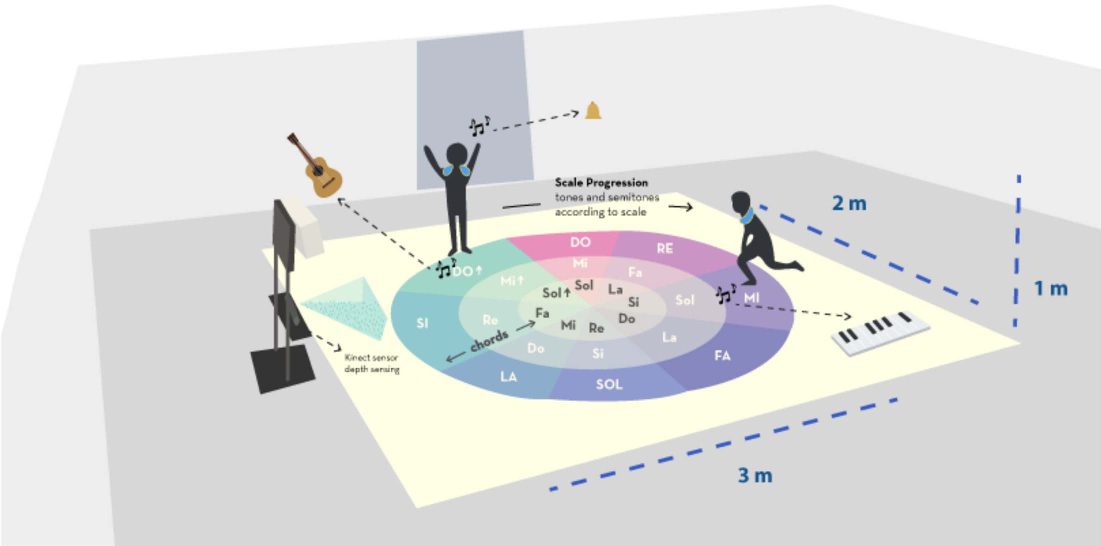


Figure 6.2: Sound environment mapping.

QoM	<->	Tempo	CI	<->	Sound Level
Low QoM	=	Slow	High CI	=	Contracted
High QoM	=	Fast	Low CI	=	Expanded

Figure 6.3: GenPlay's sound mapping based on expressive motion features (Quantity of Motion and Contraction Index).

In the second phase, it was planned to encourage collaboration between peers and introduce surprising sound elements within soundscape which could be manipulated by upper body gestures. For example, when two users come into close proximity, the sound particles start being manipulated within the soundscape. This can be perceived as a new stage in the sound interaction which encourages collaboration between pairs. It is designed to be similar to the analogy of touching hand ring bells hanging in the soundscape. Therefore the more the users are active in this state the more sound particles are manipulated. Moreover, the users are also allowed to change the volume of these sound particles based on an interaction design approach which is again suggested by Castellano's one of the expressive motion features. This feature is called the contraction index (CI) and it is associated with how the user's body expands (lower CI) or contracts (higher CI) in the surrounding space [Castellano et al., 2007]. We incorporated this design approach to allow the users to change the volume of sound particles based on their detected CI (Figure 6.3). Additionally, in order to foster social interaction through exploration within soundscape, the experience also randomly places different sound effects to the corners of the playing field which can be triggered when two participants again come into close proximity.

6.3.1.2.2 Visual Elements

In GenPlay's generative visual interactions, the visual interaction elements were based on abstract graphics. By "abstract graphics" we refer to those which do not represent a depiction of a a real-life object, but instead makes use of shapes, colours, and particle systems. The idea behind using abstract visuals was to allow children to assign their own meaning to the interaction elements which could be then a motivational drive for sharing their own experiences and fostering SIBs. For example, similar to the strategy being used in GenPlay's soundscape, first, children were able to familiarize themselves with the environment with their spatial positioning represented with two bright circles (in the shape of bubbles) around their feet within the virtual environment. When the participants came in close proximity the bubbles that mark their spatial position fused into a single one and changed color, signaling they entered a new state and that they could now collaborate. In this state, they were able to discover new surprising abstract particles in the virtual environment (Figure 6.4).

Figure 6.4: Initial interaction with GenPlay's visual interface (left) / When participants were in close proximity, their bubbles circulating their spatial position became one and changed color signaling the collaborative state to the user (right).

6.3.1.3 Data Acquisition & Data Analysis

In order to acquire data, we used the same tools described in section 3.3 Data Acquisition. such as the superhero collar for collecting psychophysiological data, video recordings for coding SIBs, questionnaire_STAIC/S and system logs (giving information towards CI and Quantity of Motion). This data was preprocessed and analyzed following the same procedures described in chapter 4 Data Analysis.

6.3.1.4 Experimental Design

This study was designed to be exploratory. The purpose of the study was to explore how full-body interaction environments can be improved in terms of fostering SIBs in children with ASC when the main interaction element is sound, especially in a classroom setting. This would allow us to develop some specific hypothesis or prediction that can be tested in future research.

GenPlay's experiment was designed in a way that offers a comparison between its soundscape and the visual interface and also the experience where both combined. In this regard, we designed the experiment based on a repeated measures design with three conditions, Condition 1 (Sound-only Interaction), Condition 2 (Visual-only Interaction), Condition 3 (Sound and Visual Interaction) with two participants going through each condition together: a child with ASC (high-functioning) and a child without ASC.

6.3.1.4.1 Participants

Testing was conducted at two different schools. We will refer to the schools as School1 and School2. School1 was a school for children with developmental disabilities and School2 was a mainstream school. Participants were two non-ASC children and eight children with ASC with ages between 7-13 years old. While their WISC IQ profiles were not provided to us, the psychologists of the schools reported to us that 5 of the ASC children were considered high-functioning, while 2 were considered mid-functioning, and 1 low-functioning.

Although the experiment was supposed to be conducted by pairs of one ASC and one non-ASC child, School1 is a special needs school, so two children with ASC had to participate together. However, the pairing of children was arranged to be distributed similarly by their level of functionality. On the other hand, the two pairs who participated from School2, being this a mainstream school, were formed as intended by one non-ASC and one ASC child.

6.3.1.4.2 Procedure

All procedures performed were aligned with the ethical approval granted by the ethics committees of the University Pompeu Fabra. Following this step, we gathered consents to participate from parents through a signed informed consent form, which detailed the goals and procedures of the project. We also gathered assents from the children on the day of the trial to make sure they were willing to participate.

Each pair of children experienced each of the conditions once for 3 minutes. Before each experimental condition, there were 1 min baseline recordings, in a standing position, looking at black screen. Before and after each condition, we administered questionnaires. The assignment of the order of the conditions was carried out randomly to counterbalance the order effect of the tasks (Figure 6.5).

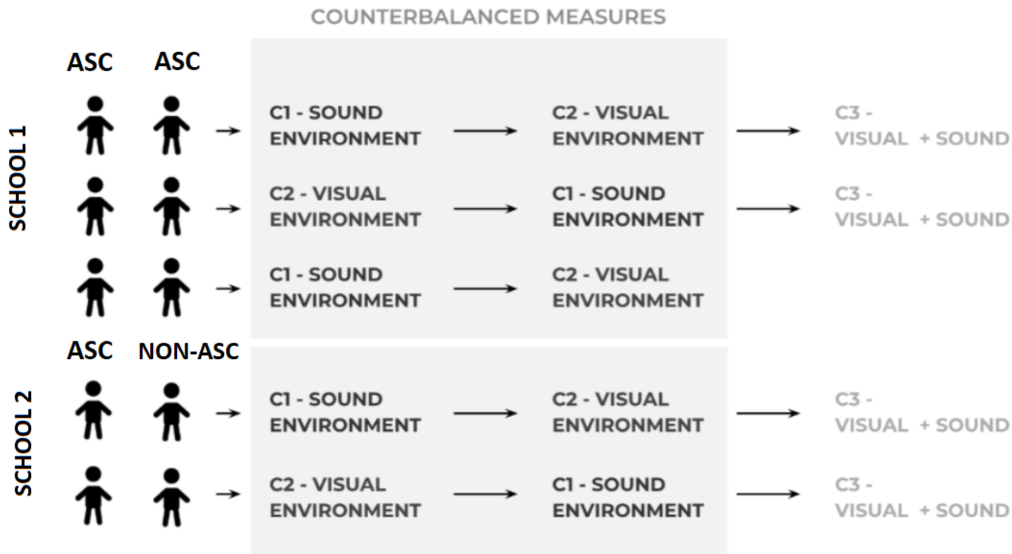


Figure 6.5: GenPlay's Experiemental Procedure & Participation.

6.4 Experiments & Results

In this section, we present the details of the statistical inference analysis to determine if there are differences in the values of measured data sources (psychophysiological, video coding, questionnaires and system log) between conditions in children with ASC. However, the data from 3 children with ASC in school 1, has been excluded due to uncompleted trials. In addition, due to the adverse lighting conditions of the experimental setting at school 1, video coding from the 3 trials was not possible. Finally, the data from psychophysiological measurements have currently been analyzed for monitoring possible movement artifacts and its feasibility to be used in a classroom setting.

6.4.1 Results

Firstly, a Friedman test was run to determine if there were differences in anxiety level (measured through STAIC/S questionnaires) after the children experienced each condition. Results show the highest STAIC/S score for the condition with sound-only interactions (C1), and the lowest score for the condition which combines both sound and visual interactions (C3) but the differences were not statistically significant. Moreover, from video coding the two experiments done at School2, it has been possible to extract and analyse relevant Social Interaction behavior data. When a Friedman test was run to determine if there were differences in the total number of SIBs during each condition, results showed that the condition with both sound and visual interactions (C3) showed the highest number of social interaction moments. However, the differences were not statistically significant. Finally, to assess the impact that the sound interactions have over the participants' expressive corporal communication, we performed a Friedman test for the CI values of the different conditions. Expressive Motion has been operationalized as the Contraction Index (CI). As a result of the Friedman test for the CI values of the different conditions we found that the Contraction Index values were significantly higher in the conditions involving sound interactions (C1-Sound and C3-Both). Children with ASC showed more expressive corporal communication in the conditions involving sound interactions (C1-Sound and C3-Both)

6.5 Discussion & Conclusion

On the question of finding differences between total number of SIBs happening in each condition, this exploratory study found that the condition with both sound and visual interactions (C3) showed the highest number of social interaction, although the differences were not significant. This was the case for reported STAIC/S scores as well. These findings might imply that the role of sound in combination with the visual elements deserves to be further investigated. In comparison to some research that regard sound as secondary to the design of visual stimulation for people with ASC [Millen et al., 2010], our initial results might suggest that music and visuals might go hand in hand and enhance more the qualities of each other. These initial findings can be considered consistent with that of [Mora Guiard, 2009] who found that visuals that are associated with music get more attention in children with ASC. Moreover, in conditions involving sound interactions, children with ASC showed more expressive corporal communication. A higher CI implies the participant has contracted and expanded their bodies a higher number of times, which can be understood as a higher amount of expressive and gestural body motion. These initial findings might be also supported by findings from some of the music therapy studies that indicate the importance of sensory aspects of a musical performance which can facilitate the need for nonverbal expression of ASC children and thereby enhance the ability to communicate [Gold et al., 2006].

6.5.1 Limitations & Future work

Taken together, GenPlay's developed sound interactions in this exploratory research are merely a starting point for the future steps of this project, which might be influenced in part by the advantages and limitations found at this stage of the design. For example, we have observed that there was a lack of awareness of the partner-triggered objects in most situations. When the system was activating these objects, the participants were not aware of the fact that the proximity of their spatial positioning were actually triggering those specific events. It is possible that the low contingency of these events was probably caused by the reduced size of the playing area, which resulted in forcing children to be in close proximity unintentionally. Moreover, it was challenging to keep good tracking of the children's movement when they were crossing infront of each other as seen from the Kinect sensor, which made the tracking system lose the participants positions for a small fraction of time and hence interrupting the game flow. A possible solution to this could be to use three new Azure Kinect systems to create a 360 view for full motion capture in the playing field. Additionally, although pychophysiological measurements were robust enough to be detected clearly in tested classroom environments, further research is needed to extract meaningful information from these measurements.

In conclusion, in terms of verifying the suitability of the method being used and determining the validity and efficacy of the experimental procedure, these pilot trials have been especially useful. The study discussed here can be seen as an early path to a future GenPlay system in which children with ASC and non-ASC can play together and benefit from it in a regular classroom setting.

Chapter 7

CONCLUSIONS

This research aimed to develop structured evaluation methods for full-body interactive technologies where children with ASC practice SIBs. Specifically, the research included implementing several data gathering methodologies and provided computational models to better understand the effectiveness of full-body interactive technologies in fostering SIBs in children with ASC. This chapter summarises the research presented, discusses the main contributions of the work and outlines limitations and the future directions of the current research.

7.1 Summary of Contributions

Contributions associated with the research questions in section 1.3.1-Research Questions will be summarized in this section.

In chapter 2, we presented a detailed literature review. We provided background information on ASC, challenges and intervention strategies. We reviewed the state of the art of ICT designed for ASC. We introduced the embodied cognition theory and our approach to contribute in the field of full-body interaction. After this discussion, we discussed specific evaluation techniques for developing social interventions for ASC. As a next step, we presented an overview of the state of the art in the computational approaches for ASC social interventions with specific focus from the perspective of multimodal evaluation. Other topics reviewed also included data sharing in ICT based ASC interventions.As it can be seen from this state of the art, no research has yet assessed the impact of an MR full-body interactive system in fostering social initiation between children with ASC and non-ASC, and compare it with a traditional intervention approach using the evaluation techniques proposed in this thesis.

In chapter 3, we discussed the materials and methods being used for developing structured evaluation methods for full-body interactive technologies where children with ASC practice SIBs. As a first step, we viewed the full-body interactive environments being tested in this thesis. We presented Lands of Fog, a large-scale full-body interactive mixed reality environment, which was designed and developed previously by FuBIntLab. Next, in order to compare the potential of Lands of Fog with a typical intervention used by therapists, we presented an adapted version of a commonly used LEGO intervention. For all of these environments, we discussed the design and development process, physical interface, game mechanics and interaction design used in detail. After this, we presented the data acquisition stage used for this work with the type of data collected, definitions to this data and the instruments being developed and used. Finally, we explained how this data is collected in the experimental design stage. The contributions of this chapter include the design and construction of a wearable which:

- allows the acquisition of accurate psychophysiological data
- in a child-friendly format,
- and is robust to variations in contact and muscular activity in the context of ambulatory full-body interaction environments.

Although this work showed a potential for submitting a utility model application, we finally decided to make this available to the research community and hence share the information on its design and construction through an open access policy. Currently, this work is under preparation for publication.

Another contribution of chapter 3 was the design of an experimental setup for comparing an MR full-body interactive system with a control condition based on construction toys (in our case LEGO bricks), which includes:

- approaches for synchronized data acquisition,
- adapted tablet based questionnaires,
- and adapted relaxation exercises.

These efforts were aligned with our first research question in terms of investigating what evaluation methods can be effective for full-body interactive systems which foster SIBs in children with ASC.

In chapter 4, we introduced the data annotation, preprocessing and feature extraction stage for all the data sources collected in this work. As the addition of psychophysiological measures to the other types of data collected from users and the system, meant that the rest of data should also be time series data and therefore timestamped to allow a proper multimodal merging of all data, we needed a high degree of time precision in video coding. In this regard the contribution of this chapter included the design and development of a novel structured video coding protocol and an adapted coding grid especially conceived for SIBs in ASC children. This contribution was aligned with our first two research questions mentioned in section 1.3.1 Research Questions and resulted in a new article which is currently under preparation for publication.

Moreover, other contributions of chapter 4 were the semi-automatic software pipeline developed for data extraction and the creation of a multimodal database which gathers all the mentioned data sources into a single resource for research. This database and the semi-automatic pipeline are already publicly available from Zenodo at DOI: 10.5281/zenodo.4557383 and from GitHub at https://github.com/batuhansayis/ASCMEOR.git .

In chapter 5, we discussed the computational approaches we have used for modelling SIBs in children with ASC during sessions in our MR environment, compared to sessions with the traditional therapy setting. These computational models are based on psychophysiological, body cues, self assessment, and game activity measurements that we presented. Specifically, we proposed three different computational models by following two different strategies.

- In the first strategy, we focused on investigating the effect generated by each social initiation moment in psychophysiological measurements and body cues separately.

- In the second strategy, we focused on investigating the observed general social state of the children with ASC in relation to all the data collected during the sessions.

These developed models contributed to addressing of all the research questions mentioned in section 1.3.1 Research Questions but with special focus on the research question 4 where we investigated the possibility of developing a full-body interactive ICT system to affect the SIBs as much as a traditional therapy setting. The first strategy, which incorporated psychophysiological measurements, has been published in the International Conference on Affective Computing and Intelligent Interaction (ACII 2019) [Sayis et al., 2019]. On the other hand, the work that incorporated body cues was undertaken, as part of a research stay, in collaboration with the Affective Intelligence & Robotics Lab at the University of Cambridge (UK) under the supervision of Prof. Hatice Gunes. This work has been published in the ACM International Conference on Multimodal Interaction, Face and Gesture Analysis for Health Informatics Workshop (ICMI 2020) [Sayis et al., 2020]. In the second strategy, the initial part of the work which includes data from the first 18 trials has been published in the Cyberpsychology, behavior, and Social Networking Journal [Crowell et al., 2019b], while the main part, which incorporates the data from all the trials and provides a computational model, has been submitted to the Virtual Reality Journal and is currently under review.

In Chapter 6, we gave information on the methodologies used to explore new interaction design techniques through sound in order to contribute to the understanding of how full-body interaction environments can be improved. We presented an exploratory study and described the design process, the development and the evaluation of a full-body interactive soundspace (GenPlay Soundscape). With this additional research, we contributed to the research question 3 in section 1.3.1 Research Questions and we demonstrated that full-body interaction is a suitable approximation for developing systems to engage users and motivate them to create SIBs with the help of sound. This study was part of a larger project called Generative Playscape (GenPlay) that aims to supply novel resources to the teachers of integrated classroom settings which would serve as a support for increasing inclusion during playtime activities. Our work associated to this project was chosen among the top 12 entries in the Student Design Competition of the 2018 CHI Conference on Human Factors in Computing Systems (CHI 2018), among a pool of entries from international research teams, and was invited to be presented at the conference and published in the Extended Abstracts of the conference [Crowell et al., 2018]

7.2 Results & Discussion

This section summarises the key findings from each of the main contributions and results, their impact and their relation to the ability of full-body interactive technologies to foster SIBs between children with ASC and non-ASC, and a detailed analysis of the specific interaction design concepts designed for Lands of Fog.

7.2.1 A new wearable to understand the ANS in Autism Research

Wearable sensors are a promising research direction for monitoring psychophysiological states in individuals with ASC [Picard, 2009]. However, when wearable technologies have to be applied in ASC research, particular attention should be paid to attributes such as comfort and unobtrusiveness [Koo et al., 2020, Taj-Eldin et al., 2018]. Moreover, physiological signals are prone to contamination by motion artifacts [Boucsein, 2012]. Particular consideration should be paid, especially in full-body interaction environments, to using both ECG and EDA, since they are susceptible to motion. In this regard, we designed and developed a wearable specifically for children with ASC, optimizing the monitoring of pschophysiological signals in full-body interaction environments. The results from the verification and validation studies were promising. Moreover, in the experimental trials it proved to be robust as well as a fun and acceptable experience for children (n = 72). Specifically, the shoulder electrode placement compared to other parts of the commonly used EDA measurement locations (i.e. finger palm) proved useful in expanding the knowledge around the usage of similar EDA electrode positioning in full-body interaction environments. Moreover, as the wearable form factor design that we chose as a "super hero's cape" was acceptable by children, using such child-friendly design can be a way to approach introducing new type of wearable in Autism Research. Finally, as this work was validated in ecologically valid full-body interactive settings (including a classroom environment), it has a potential to fit in the broader scope of understanding the ANS in Autism Research, and in many other conditions.

7.2.2 A novel way of coding interpersonal social behaviors in HCI interventions in children with ASC

When coding behaviors during interventions using an interactive system, much richer results can be achieved if the data from video coding is integrated with other data; e.g. system data (log files of system decisions and actions) or psychophysiological data. These are all time series data that keep track of what is going on with the participants at every instant. Relating this data would allow the generation of a shared multimodal representation that could be used in understanding the deeper details of specific behaviors from multiple points of view. This implies that video coding strategies applied on such interactive systems should be flexible enough (on a microanalytic basis) to allow for their integration with different modalities. In this regard, as we saw in section 2.6.1 Video coding, no existing observation scheme was found that allowed reliable time-stamped event coding. We did not find any either that accounted for full-body interactive settings where the system plays a role in fostering SIBs towards or between the children with ASC. Therefore, we developed an adapted observation protocol through an iterative development process. Our results showed that with the developed protocol it was possible to reach acceptable inter-rater agreement scores. Moreover, it was shown that this protocol is flexible enough to code the range of behaviors presented both in a technological intervention and a non-technological intervention for coding SIBs in ASC and non-ASC. This video coding protocol, with its step by step flow chart, and coding grid, provides a novel way of coding interpersonal social behaviors in HCI interventions in children with ASC.

7.2.3 Computational Models to investigate the effectiveness of Full-body interaction environments in fostering SIBs in children with ASC

Full-body Interaction experiences, based on technological configurations within the realm of Mixed Reality started playing an important role in addressing the issue of encouraging socialization behaviors in children with Autism. With the help of a design of a comparative experimental setup and implementation of new data gathering and evaluation techniques we were able to show the effectiveness of our MR system compared to a traditional intervention. The results showed that our MR system was as effective as the traditional therapy setting based on LEGO bricks, but with an added advantage of being more flexible.The self-contained nature of the MR system was able to detect and respond to user activity based on accomplishments, and could adapt the game's response accordingly. This was also perceived by the psychologist in our team, who said that the active nature of the MR system was unique in that the responses were clear and immediate in a way that he could not replicate in his own therapy practices.

When this effectiveness was investigated in terms of the observed general social state of the children with the help of complex statistical inference and computational modeling techniques, it was seen that our system was capable of fostering SIBs in children with ASC with similar success as the LEGO setting (section 5.4.3 Experiments & Results). These results imply that our system is capable of fostering SIBs in children with ASC with similar success as the traditional strategies, such as playing with LEGO bricks. This is important since there is no human intervention in motivating the children in our system, whereas traditional strategies depend on the abilities of the therapist to foster SIBs in children. This does not mean that we propose an unsupervised system that does without experts. On the contrary, our system can, on the one hand, free the experts from having to mediate the session and allows them to focus on observing the evolution of the child within the intervention program. Moreover, analyzing the questionnaires, we saw that no significant difference existed between the two conditions regarding the reported level of knowledge of the play peer, and the desire to know more about the peer after playing. In this regard, the promotion of collaboration in the MR environment is probably helping ASC children to see their peers as valuable play partners, in a similar way to how they see peers in the LEGO condition. Given the challenges in social integration and maintaining social relationships in children with ASC [Strain, 1983, Gupta et al., 2014], creating scenarios where users see each other as valuable partners could help in developing and maintaining friendship. Lands of Fog experience offers the possibility of having a broad diversity of scenarios and can therefore provide even more possibilities to create collaboration situations, compared to a static unchanging approach such as the LEGO bricks are. In terms of SIBs' relation to the internal states of the ASC participants, results suggest similar levels of anxiety, similar levels of social functioning through HRV, and similar levels of engagement through EDA-based measurements in both conditions. Moreover, our results showed that ASC participants' desire to know more about their play partner after playing were coupled with their EDA measurements in Lands of Fog but not in LEGO (section 5.4.3 Experiments & Results). This might suggest that Lands of Fog is probably creating an adequate engaging environment for getting to know the peer and hence fostering generation of SIBs. In terms of SIBs relation to the events generated and sensed by the system, basing our design on the encouraged collaboration approach proved itself to be helpful in fostering SIBs (section 5.4.3 Experiments & Results). One of the important findings was that we can keep up engagement in children by providing sufficient interaction potential, using game elements and mechanics such as hunting fireflies and manipulating props collaboratively. Also that we allow the children to step back from collaborative play and go back to solo playing at any one time.

On the other hand, in order to investigate this effectiveness in terms of the characteristics of each social initiation behavior, we proposed computational models which are based on psychophysiological measurements and on body cues separately. The proposed computational model based on psychophysiological measurements helped us see that certain HRV and EDA features might be associated with social functioning, anxiety, or physical activity during Lands of Fog and LEGO conditions. Moreover, it was shown that these features could have an effect on finding the differences between the social initiation taking place in the MR condition and that taking place in the LEGO condition (section 5.2.3.3 Results). This shows that our MR system has specific properties, compared to a traditional construction-based intervention, which potentially provides a new interesting context to interventions in ASC. It could be argued that the differences found between social initiation moments were due to diversity of activities presented in each condition which can confound the social initiation related arousal activity. However, here with this study we also highlight the effect of some of the confounding variables. For example, we are presenting the effect of the features which might be more related to physical activity and the features which might be related to social functioning or anxiety. The diversity of the activities which might be related to these findings might be acknowledged as the part of the interesting context specified but this requires a further investigation. In this regard, this study can pave the way to better understand whether practicing socialization in an MR environment catered towards individuals with ASC could be a way to reduce anxiety levels while simultaneously forming collaborative behavioral patterns.

The proposed computational models, which are based on body cues, allowed us to see that the combination of certain body features help us recognize the type of social initiation moment (ASC or non-ASC) in both conditions showing that ASC and non-ASC possibly present different nonverbal body behaviors in social initiation moments (section 5.3.3 Experiments and Results).Finding this difference not significant between conditions might be interpreted as the MR system encouraging similar nonverbal behaviors in children. Moreover, the similarity of nonverbal behaviors between ASC and non-ASC children seems to be greater in the MR condition than in the LEGO condition, as the performance scores in the MR setting are lower as compared to the LEGO setting.In this regard, when a further investigation was run through manual video analysis of the body cues, we observed that during the social initiation moments, when exploration is taking place in the MR system play area, ASC children had a tendency to follow the non-ASC children and they would keep a close proximity. Data from several studies suggests that the ASC child usually has less chance to take a leading role in activities with mixed dyads of friendship. Therefore, settings that mediate communication between individuals should encourage balanced interactions to ensure that children with ASC have an equal opportunity in leading the interactions and the game [Bauminger et al., 2008]. As we were able to observe ASC initiation taking place at these close proximity moments, it can be said that Lands of Fog might be providing balanced opportunity for ASC children to take the lead. However, the ASC child is still following the lead of the non-ASC child and therefore we must continue research on the interaction design to foster leadership of the ASC child during the experience.

It should be also noted that the proposed computational models that allowed us to observe the mentioned findings have a potential to be used in different scenarios. The proposed computational model that is based on psychophysiological measurements helps in recognising social initiation induced arousal in ASC across different conditions. Moreover, our approach based on individual and interpersonal body cues, helps in recognising social initiation nonverbal behavior differences between ASC and non-ASC in a technological and non-technological setting, also providing a comparison measure across settings. The proposed computational models based on multimodal data help in recognising contextualized levels of social initiations in children with ASC in a technological and non-technological setting.

7.2.4 One of the first multimodal databases recording ASC and non-ASC children's interpersonal social behaviors

It is worth nothing that ICT studies for ASC do not appear motivated to share their data to other researchers. Furthermore, the datasets currently available in our context are not focusing on social interaction between ASC and non-ASC children. In this regard, public release of ASCMEOR database provides a unique opportunity for furthering our research in a data-driven collaboration with other researchers. The ASCMEOR database can facilitate further development of computational models of SIBs. ASCMEOR database has already been used at Affective Intelligence & Robotics Lab in the context of analyzing individual and interpersonal body cues in children with ASC and non-ASC.

7.2.5 An approximation for developing sound based full-body interactive MR systems to foster SIBs

After our research in interaction design techniques using sound in the context of full-body interactive MR environments, we have found that this combination of sound and visuals deserves further attention. We must address the already existing knowledge in this field and take advantage of the potential of music and visuals going hand in hand and enhance each other's qualities. This could also support the idea that the richer multimodal experiences can better attract children with ASC in creating social interactions. Moreover, in conditions involving sound interactions, children with ASC showed more expressive corporal communication. A higher CI implies the participant has contracted and expanded their bodies a higher number of times, which can be understood as a higher amount of expressive and gestural body motion. These initial findings might be also supported by findings from some of the music therapy studies that indicate the importance of sensory aspects of a musical performance which can facilitate the need for nonverbal expression of ASC children and thereby enhance the ability to communicate. However, GenPlay's developed sound interactions in this exploratory research were merely a starting point. It opened the door to starting a new research project which is now starting in the FuBIntLab.

7.3 Limitations & Future Work

The knowledge gained from this research has contributed to the development of structured monitoring and evaluation methods for full-body interactive technologies where children with ASC practice SIBs. Specifically, this study is a step towards ultimately implementing several data gathering methodologies and developing computational models to better understand the effectiveness of full-body interactive technologies in fostering SIBs in children with ASC. Nonehteless, while conducting our research, a number of limitations have become apparent. Among the main issues of this study were the paucity of recruiting child users and the extended time frame needed for developing new evaluation tools.

7.3.1 Logistical Limitations

One of the purposes of this research was to bring ASC children together with the non-ASC children creating an ecologically valid environment in the sense that it defines a similar situation to that of meeting other children in a school playground or a public park. On the one hand, this would also help the Autism community to better cope with social integration challenges and, on the other hand, make non-ASC children see autism in a more natural way and help better integrate ASC cildren in society.

Given this need of having one ASC and one non-ASC child interacting together during an experimental trial, and given the duration of the trials (which was over 1 hour), we had to arrange specific visits of the two children, at concrete scheduled times, to our lab for each trial session. Because of this, we could not perform (as has been done in many other projects of the FuBIntLab) the recruitment of full school class groups and have them all come to our lab in one batch. This would have not allowed us to have all the 25+ children go through the trials in one morning visit, for example. Even less when we would have needed one group of ASC and one of non-ASC children at the same time to have the mixed trials.

This difficult challenge of finding participants at specified time slots caused important delays in the evolution of the research. Considering that children between 8-12 years old had to be accompanied by their parents (or legal representatives) during the experimental trials, their participation depended completely on the willingness of the parents to participate in our study.

Thanks to our collaboration with Hospital Sant Joan de Déu, we had good access to a number of autism institutions which supposed an important pool of potential ASC participants. These were were individually contacted through a time-consuming process of phone interviews with the parents to double check that their children met the set of inclusion criteria specified in our experimental design (see section 3.4.1 Demography). Despite these difficulties, parents of ASC children are in general very motivated to help their children in any possible way and, hence, were positive about participating in our research.

However, reaching to non-ASC participants was much more challenging as we had to build our user base from scratch. This required formalizing several dissemination and communication strategies including: distributing flyers at a shopping mall close to our university (advertising a large-scale video game that was free to play); launching a social media campaign through Facebook; having meetings with civic centers and schools; etc. These efforts helped us reach a number of potential participants but many of the initially contacted parents did not sign up for the trial after the first contact. The issue here is that parents of non-ASC children are not so motivated to participate in the study since their week is already quite full with activities. Despite they may have strong social values and potentially wish to support our research, their family logistics clashed with having to come to our Lab. This created a strong slowing down of our trials and put us in a difficult situation.

Towards the end of this research, we had the chance to start a collaboration with the FAPAC, the Federation of Asscociations of Student's Mothers and Fathers in Catalonia. This started to yield very fruitful results, but only at a moment in which the research had to be closed. Nonetheless, this is a very positive relationship for future extensions of this research.

7.3.2 Methodological limitations

As the creation of a new coding protocol allowed proper multimodal merging of all the data, this stage was crucial to get progress in our research. However, this process took longer than expected (approximately one year and a half) since creating the finalized version of the protocol depended upon reaching an acceptable interrater agreement which was directly dependent on the number of videos available at that moment (associated with the logistic limitations mentioned in section 7.3.1 Logistical Limitations).

The logistic limitations also caused data to be collected in a discontinuous manner over the course of this thesis work (sometimes we were able to run just 1 trial a month). As a result, data analysis and computational modeling stages had to be conducted in a flexible iterative manner till the initial necessary amount of data was collected, limiting the time frame of thorough analysis of the results.

Research on body cues opened a dimension in our study where the role of interpersonal communication between ASC and non-ASC children (not just the ASC activity) started being investigated in terms of fostering SIBs in full-body interaction environments. Through this initiative we started coding non-ASC SIBs which could be only incorporated into data analysis in the final stages of this thesis work. Moreover, research in body cues also included investigating the potential of facial expressions. However, the camera positioning in this experimental setup did not allow to detect facial regions effectively.

7.3.3 Future Work

The results of this research support the idea that our system is capable of fostering SIBs in children with ASC with similar success as the LEGO condition, with an added advantage of being more flexible. This finding reported here shed new light on developing a tool that is mediating, guiding, and supporting the progress of the children in terms of practicing SIBs; and providing structure and assistance to therapists. Moreover, a further longitudinal study could assess the long-term effects of this tool, especially when it is used in place of a passive social skills training program that needs to be mediated by a therapist.

The results of this study also indicate that SIBs are related to some of the internal state measurements. Nevertheless, physiological measures are susceptible to physical activity. For future research, physiological measurements may be standardized with activity levels recognized from ACC sensors or with the quantity of body motion information, which can be tracked with special algorithms through camera systems. Additionally, as a next step, physiological measurements can be coupled with computer vision based internal state evaluation (e.g face and body gesture recognition) for more accurate evaluation. Furthermore, since our wearable allows hosting different types of sensors, we can incorporate different modalities to our data analysis pipeline such as a dedicated michrophone for detailed natural language processing, respiration rate, and a temperature sensor. Implementation of such multimodal objective evaluation tools can pave the way for carefully designing and assessing the novel tools that can be utilised by the ASC therapists and caregivers.

Finally, the results of this investigation also show that SIBs are related to some of the game activity measurements. The results have provided very useful insight on which game mechanics were useful and which were not. With this information we can improve the MR experience for the children, both by improving the current game mechanics, as well as designing new ones that can appeal to a broader range of children in the ASC. In this regard, as a continuation of Land of Fog, we have submitted a grant proposal to the Spanish Ministry of Economy and Business to expand the current system, to understand how we can help children with ASC adopt Prosocial behaviors through full-body interaction using also sophisticated sound interaction elements in addition to visual elements.

Bibliography

[Adams, 2013] Adams, G. F. (2013). Video Tape Analysis Studies BT - Encyclopedia of Autism Spectrum Disorders. pages 3275–3281. Springer New York, New York, NY. 18

[Agarwal et al., 2012] Agarwal, R., Sampath, H., and Indurkhya, B. (2012). AUTINECT-Helping children with autism acquire social skills through virtual peers. 17

[Ajili et al., 2018] Ajili, I., Mallem, M., and Didier, J.-Y. (2018). Relevant LMA Features for Human Motion Recognition. 80

[Alcorn et al., 2011] Alcorn, A., Pain, H., Rajendran, G., Smith, T., Lemon, O., Porayska-Pomsta, K., Foster, M. E., Avramides, K., Frauenberger, C., and Bernardini, S. (2011). Social communication between virtual characters and children with autism. In *international conference on artificial intelligence in education*, pages 7–14. Springer. 2, 14

[Alexander et al., 1995] Alexander, J. F., Newell, R. M., Robbins, M. S., and Turner, C. W. (1995). Observational coding in family therapy process research. *Journal of Family Psychology*, 9(4):355. 18, 19

[Antle et al., 2009] Antle, A. N., Corness, G., and Droumeva, M. (2009). What the body knows: Exploring the benefits of embodied metaphors in hybrid physical digital environments. *Interacting with Computers*, 21(1-2):66–75. 2, 17

[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. 23

[Ardalan et al., 2019] Ardalan, A., Assadi, A. H., Surgent, O. J., and Travers, B. G. (2019). Whole-body movement during videogame play distinguishes youth with autism from youth with typical development. *Scientific reports*, 9(1):1–11. 23

[Association, 2013] Association, A. P. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Washington, DC, 5 edition. 1, 12

[Attwood, 2014] Attwood, T. (2014). Frameworks for Behavioral Interventions to Improve peer relationships and the management of emotions in Asperger syndrome. *ebook*, page 239. 48

[Autismspeaks.org,] Autismspeaks.org. Parent's Guide to Blood Draws for Children with Autism A Parent's Guide To Blood Draws for Children with Autism. xvi, 48, 150

[Azuma, 1997] Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, 6(4):355–385. 15

[Bai et al., 2016] Bai, J., Di, C., Xiao, L., Evenson, K. R., LaCroix, A. Z., Crainiceanu, C. M., and Buchner, D. M. (2016). An activity index for raw accelerometry data and its comparison with other activity metrics. *PloS one*, 11(8):e0160644. 64

[Bai et al., 2013] Bai, Z., Blackwell, A. F., and Coulouris, G. (2013). Through the looking glass: Pretend play for children with autism. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 49–58. IEEE. 15

[Bakker et al., 2008] Bakker, S., Markopoulos, P., and Kort, Y. D. (2008). OPOS: an observation scheme for evaluating head-up play. *... of the 5th Nordic conference on ...*, pages 18–22. 20

[Bannach et al., 2009] Bannach, D., Amft, O., and Lukowicz, P. (2009). Automatic event-based synchronization of multimodal data streams from wearable and ambient sensors. In *European Conference on Smart Sensing and Context*, pages 135–148. Springer. 49

[Baranek, 2002] Baranek, G. T. (2002). Efficacy of sensory and motor interventions for children with autism. *Journal of autism and developmental disorders*, 32(5):397–422. 17

[Barber, 2013] Barber, M. (2013). United States Publication. *School Science and Mathematics*, 25(2):181–181. 39, 41

[Baron-Cohen, 2008] Baron-Cohen, S. (2008). *Autism and Asperger Syndrome*. Oxford University Press. 1, 2, 13

[Bartoli et al., 2013] Bartoli, L., Corradi, C., Garzotto, F., and Valoriani, M. (2013). Exploring motion-based touchless games for autistic children's learning. In *Proceedings of the 12th International Conference on Interaction Design and Children - IDC '13*, pages 102–111, New York, New York, USA. ACM Press. 16

[Bauminger, 2002] Bauminger, N. (2002). The facilitation of social-emotional understanding and social interaction in high-functioning children with autism: Intervention outcomes. *Journal of autism and developmental disorders*, 32(4):283–298. 19, 56

[Bauminger et al., 2008] Bauminger, N., Solomon, M., Aviezer, A., Heung, K., Brown, J., and Rogers, S. J. (2008). Friendship in high-functioning children with autism spectrum disorder: Mixed and non-mixed dyads. *Journal of Autism and Developmental Disorders*, 38(7):1211–1229. 86, 121

[Bauminger et al., 2010] Bauminger, N., Solomon, M., and Rogers, S. J. (2010). Externalizing and internalizing behaviors in ASD. *Autism research : official journal of the International Society for Autism Research*, 3(3):101–12. 35

[Bellini, 2006] Bellini, S. (2006). The development of social anxiety in adolescents with autism spectrum disorders. *Focus on autism and other developmental disabilities*, 21(3):138–145. 13

[Ben-Sasson et al., 2013] Ben-Sasson, A., Lamash, L., and Gal, E. (2013). To enforce or not to enforce? The use of collaborative interfaces to promote social skills in children with high functioning autism spectrum disorder. *Autism*, 17(5):608–622. 35

[Benedek and Kaernbach, 2010] Benedek, M. and Kaernbach, C. (2010). Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology*, 47(4):647–658. 63, 64, 74

[Benevides and Lane, 2013] Benevides, T. W. and Lane, S. J. (2013). A Review of Cardiac Autonomic Measures: Considerations for Examination of Physiological Response in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 45(2):560–575. 22

[Bent et al., 2020] Bent, B., Goldstein, B. A., Kibbe, W. A., and Dunn, J. P. (2020). Investigating sources of inaccuracy in wearable optical heart rate sensors. *npj Digital Medicine*, 3(1):1–9. 38

[Bernard-Opitz et al., 2001] Bernard-Opitz, V., Sriram, N., and Nakhoda-Sapuan, S. (2001). Enhancing social problem solving in children with autism and normal children through computer-assisted instruction. *Journal of autism and developmental disorders*, 31(4):377–84. 14

[Bernardini et al., 2014] Bernardini, S., Porayska-Pomsta, K., and Smith, T. J. (2014). ECHOES: An intelligent serious game for fostering social communication in children with autism. *Information Sciences*, 264:41–60. 17

[Bernhardt and Robinson, 2007] Bernhardt, D. and Robinson, P. (2007). Detecting affect from non-stylised body motions. In *International conference on affective computing and intelligent interaction*, pages 59–70. Springer. 80

[Berntson et al., 2017] Berntson, G. G., Quigley, K. S., Norman, G. J., and Lozano, D. L. (2017). Cardiovascular psychophysiology. 21

[Berthoz and Hill, 2005] Berthoz, S. and Hill, E. L. (2005). The validity of using self-reports to assess emotion regulation abilities in adults with autism spectrum disorder. *European psychiatry*, 20(3):291–298. 98

[Betella and Verschure, 2016] Betella, A. and Verschure, P. F. M. J. (2016). The Affective Slider: A Digital Self-Assessment Scale for the Measurement of Human Emotions. *PLOS ONE*, 11(2):e0148037. 47

[Betsworth et al., 2014] Betsworth, L., Bowen, H., Robinson, S., and Jones, M. (2014). Performative technologies for heritage site regeneration. *Personal and Ubiquitous Computing*, 18(7):1631–1650. 16

[Bhattacharya et al., 2015] Bhattacharya, A., Gelsomini, M., Pérez-Fuster, P., Abowd, G. D., and Rozga, A. (2015). Designing motion-based activities to engage students with autism in classroom settings. In *Proceedings of the 14th International Conference on Interaction Design and Children - IDC '15*, pages 69–78, New York, New York, USA. ACM Press. 17

[Bianchi-Berthouze, 2013] Bianchi-Berthouze, N. (2013). Understanding the role of body movement in player engagement. *Human–Computer Interaction*, 28(1):40–75. 20

[Blanche et al., 2012] Blanche, E. I., Reinoso, G., Chang, M. C., and Bodison, S. (2012). Proprioceptive processing difficulties among children with autism spectrum disorders and developmental disabilities. *The American journal of occupational therapy : official publication of the American Occupational Therapy Association*, 66(5):621–624. 24

[Boettger et al., 2010] Boettger, S., Puta, C., Yeragani, V. K., Donath, L., Mueller, H.-J., Gabriel, H. H., and Baer, K.-J. (2010). Heart rate variability, QT variability, and electrodermal activity during exercise. *Med Sci Sports Exerc*, 42(3):443–448. 99

[Bone et al., 2016] Bone, D., Chaspari, T., and Narayanan, S. (2016). Chapter 15 Behavioral signal processing and autism. *Autism Imaging and Devices*, pages 319–344. 1, 25

[Bone et al., 2015] Bone, D., Goodwin, M. S., Black, M. P., Lee, C.-C., Audhkhasi, K., and Narayanan, S. (2015). Applying machine learning to facilitate autism diagnostics: pitfalls and promises. *Journal of autism and developmental disorders*, 45(5):1121–1136. 1

[Borghi and Cimatti, 2010] Borghi, A. M. and Cimatti, F. (2010). Embodied cognition and beyond: acting and sensing the body. *Neuropsychologia*, 48(3):763–73. 16

[Boucsein, 2012] Boucsein, W. (2012). *Electrodermal activity*. Springer Science & Business Media. 21, 38, 99, 118

[Braithwaite et al., 2013] Braithwaite, J. J., Watson, D. G., Jones, R., and Rowe, M. (2013). A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. *Psychophysiology*, 49(1):1017–1034. 63

[Brisson et al., 2012] Brisson, J., Warreyn, P., Serres, J., Foussier, S., and Adrien-Louis, J. (2012). Motor anticipation failure in infants with autism: a retrospective analysis of feeding situations. *Autism*, 16(4):420–429. 23

[Brown and Murray, 2001] Brown, J. and Murray, D. (2001). Strategies for enhancing play skills for children with autism spectrum disorder. *Education and Training in Mental Retardation and Developmental Disabilities*, pages 312–317. 14

[Buchmann et al., 2004] Buchmann, V., Violich, S., Billinghurst, M., and Cockburn, A. (2004). FingARtips: gesture based direct manipulation in Augmented Reality. In *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, pages 212–221. 15

[Cacioppo et al., 2007] Cacioppo, J. T., Tassinary, L. G., and Berntson, G. (2007). *Handbook of psychophysiology*. Cambridge university press. 21, 76

[Cao and Can,] Cao, Y. and Can, O. Leveraging Convolutional Pose Machines for Fast and Accurate Head Pose Estimation. 65

[Cao et al., 2018] Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., and Sheikh, Y. (2018). OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. *arXiv preprint arXiv:1812.08008*. 65

[Carlson and Grotevant, 1987] Carlson, C. I. and Grotevant, H. D. (1987). A comparative review of family rating scales: Guidelines for clinicians and researchers. *Journal of Family Psychology*, 1(1):23. 19

[Carr, 1977] Carr, E. G. (1977). The motivation of self-injurious behavior: a review of some hypotheses. *Psychological bulletin*, 84(4):800–16. 2, 13

[Casas et al., 2012] Casas, X., Herrera, G., Coma, I., and Fernandez, M. (2012). A KINECT-BASED AUGMENTED REALITY SYSTEM FOR INDIVIDUALS WITH AUTISM SPECTRUM DISORDERS. In *Proceedings of the International Conference on Computer Graphics Theory and Applications*, pages 440–446. SciTePress - Science and and Technology Publications. 2, 17

[Casenhiser et al., 2013] Casenhiser, D. M., Shanker, S. G., and Stieben, J. (2013). Learning through interaction in children with autism: preliminary data from asocial-communication-based intervention. *Autism : the international journal of research and practice*, 17(2):220–41. 13

[Castellano et al., 2007] Castellano, G., Bresin, R., Camurri, A., and Volpe, G. (2007). User-centered control of audio and visual expressive feedback by full-body movements. *Affective Computing and . . .* , pages 501–510. 108, 109

[Cebula, 2012] Cebula, K. R. (2012). Applied behavior analysis programs for autism: Sibling psychosocial adjustment during and following intervention use. *Journal of autism and developmental disorders*, 42(5):847–862. 13

[Chen and Bernard-Opitz, 1993] Chen, S. H. and Bernard-Opitz, V. (1993). Comparison of personal and computer-assisted instruction for children with autism. *Mental retardation*, 31(6):368–76. 14

[Chen, 2012] Chen, W. (2012). Multitouch Tabletop Technology for People with Autism Spectrum Disorder: A Review of the Literature. *Procedia Computer Science*, 14(1877):198–207. 16

[Chiarotti and Venerosi, 2020] Chiarotti, F. and Venerosi, A. (2020). Epidemiology of autism spectrum disorders: a review of worldwide prevalence estimates since 2014. *Brain sciences*, 10(5):274. 1

[Clearinghouse, 2013] Clearinghouse, W. W. (2013). US Department of Education, Institute of Education Sciences, What Works Clearinghouse. *Retrieved from What Works Clearinghouse: Publications & Products: http://whatworks. ed. gov*. 63

[Colton et al., 2009] Colton, M. B., Ricks, D. J., Goodrich, M. A., Dariush, B., Fujimura, K., and Fujiki, M. (2009). Toward therapist-in-the-loop assistive robotics for children with autism and specific language impairment. *autism*, 24:25. 23

[Cook et al., 2013] Cook, J. L., Blakemore, S.-J., and Press, C. (2013). Atypical basic movement kinematics in autism spectrum conditions. *Brain*, 136(9):2816–2824. 78

[CoolBlades, 2021] CoolBlades (2021). Hairline Economy Cutting Collar - CoolBlades Professional Hair & Beauty Supplies & Salon Equipment Wholesalers. xv, 41

[Coppola et al., 2016] Coppola, C., Faria, D. R., Nunes, U., and Bellotto, N. (2016). Social activity recognition based on probabilistic merging of skeleton features with proximity priors from rgb-d data. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5055–5061. IEEE. 24, 65

[Costa et al., 2001] Costa, M., Dinsbach, W., Manstead, A. S. R., and Bitti, P. E. R. (2001). Social presence, embarrassment, and nonverbal behavior. *Journal of Nonverbal Behavior*, 25(4):225–240. 85

[Critchley, 2002] Critchley, H. D. (2002). Electrodermal responses: What happens in the brain. *Neuroscientist*, 8(2):132–142. 22

[Crowell et al., 2019a] Crowell, C., Mora-Guiard, J., and Pares, N. (2019a). Structuring collaboration: Multi-user full-body interaction environments for children with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders*, 58(November 2018):96–110. 35

[Crowell et al., 2019b] Crowell, C., Sayis, B., Benitez, J. P., and Pares, N. (2019b). Mixed Reality, Full-Body Interactive Experience to Encourage Social Initiation for Autism: Comparison with a Control Nondigital Intervention. *Cyberpsychology, Behavior, and Social Networking*, page cyber.2019.0115. 117

[Crowell et al., 2018] Crowell, C., Sayis, B., Bravo, A., and Paramithiotti, A. (2018). GenPlay: Generative Playscape. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pages 1–6, New York, New York, USA. ACM Press. 117

[Dalsgaard and Dindler, 2014] Dalsgaard, P. and Dindler, C. (2014). Between theory and practice: bridging concepts in HCI research. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 1635–1644. 34

[Dautenhahn et al., 2002] Dautenhahn, K., Ogden, B., and Quick, T. (2002). From embodied to socially embedded agents - Implications for interaction-aware robots. *Cognitive Systems Research*, 3(3):397–428. 14

[Davis et al., 2010] Davis, M., Dautenhahn, K., Powell, S. D., and Nehaniv, C. L. (2010). Guidelines for researchers and practitioners designing software and software trials for children with autism. *Journal of Assistive Technologies*, 4(1):38–48. 34

[Dawson et al., 2018] Dawson, G., Campbell, K., Hashemi, J., Lippmann, S. J., Smith, V., Carpenter, K., Egger, H., Espinosa, S., Vermeer, S., and Baker, J. (2018). Atypical postural control can be detected via computer vision analysis in toddlers with autism spectrum disorder. *Scientific reports*, 8(1):1–7. 23

[Dawson et al., 2012] Dawson, G., Jones, E. J. H., Merkle, K., Venema, K., Lowy, R., Faja, S., Kamara, D., Murias, M., Greenson, J., Winter, J., Smith, M., Rogers, S. J., and Webb, S. J. (2012). Early behavioral intervention is associated with normalized brain activity in young children with autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51(11):1150–9. 13

[de Belen et al., 2020] de Belen, R. A. J., Bednarz, T., Sowmya, A., and Del Favero, D. (2020). Computer vision in autism spectrum disorder research: a systematic review of published studies from 2009 to 2019. *Translational psychiatry*, 10(1):1–20. 23, 25

[De Jaegher, 2013] De Jaegher, H. (2013). Embodiment and sense-making in autism. *Frontiers in integrative neuroscience*, 7:15. 17

[De Jaegher et al., 2013] De Jaegher, H., Jaegher, H. D., Leary, M. R., and Practice, P. (2013). Embodiment and sense-making in autism. *Frontiers in integrative neuroscience*, 7(March):15. 78

[Denny et al., 2018] Denny, E., Folkes, W., Ghattas, I., Kaufmann, H., Williams, H., and Potvin, M.-C. (2018). A Systematic Review of the Efficacy of Weighted Vests and Blankets on People with ASD or ADHD. 39, 40

[Devita-Raeburn, 2016] Devita-Raeburn, E. (2016). The controversy over autism's most common therapy — Spectrum — Autism Research News. 13

[D'Mello and Kory, 2012] D'Mello, S. and Kory, J. (2012). Consistent but modest: a meta-analysis on unimodal and multimodal affect detection accuracies from 30 studies. In *Proceedings of the 14th ACM international conference on Multimodal interaction*, pages 31–38. 101

[Donnellan et al., 2013] Donnellan, A. M., Hill, D. A., and Leary, M. R. (2013). Rethinking autism: implications of sensory and movement differences for understanding and support. *Frontiers in integrative neuroscience*, 6:124. 17

[Downey and Rapport, 2012] Downey, R. and Rapport, M. J. K. (2012). Motor Activity in Children With Autism : A Review of. 78

[Dumas et al., 2009] Dumas, B., Lalanne, D., and Oviatt, S. (2009). Multimodal interfaces: A survey of principles, models and frameworks. In *Human machine interaction*, pages 3–26. Springer. 2, 25

[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*. 14

[Eapen et al., 2013] Eapen, V., Crnčec, R., and Walter, A. (2013). Clinical outcomes of an early intervention program for preschool children with Autism Spectrum Disorder in a community group setting. *BMC pediatrics*, 13(1):3. 13

[El-Ghoroury and Romanczyk, 1999] El-Ghoroury, N. H. and Romanczyk, R. G. (1999). Play interactions of family members towards children with autism. *Journal of autism and developmental disorders*, 29(3):249–258. 86

[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., Yasamy, M. T., and Fombonne, E. (2012). Global Prevalence of Autism and Other Pervasive Developmental Disorders. *Autism Research*, 5(3):160–179. 1

[Erickson et al., 2012] Erickson, L. C., Scott-Van Zeeland, A. A., Hamilton, G., Lincoln, A., and Golomb, B. A. (2012). Brief Report: Approaches to 31 P-MRS in Awake, Non-Sedated Children with and without Autism Spectrum Disorder. *Journal of autism and developmental disorders*, 42(6):1120–1126. 40

[Esteban et al., 2017] Esteban, P. G., Baxter, P., Belpaeme, T., Billing, E., Cai, H., Cao, H.-L., Coeckelbergh, M., Costescu, C., David, D., and De Beir, A. (2017). How to build a supervised autonomous system for robot-enhanced therapy for children with autism spectrum disorder. *Paladyn, Journal of Behavioral Robotics*, 8(1):18–38. 23, 25

[Estes et al., 2014] Estes, A., Vismara, L., Mercado, C., Fitzpatrick, A., Elder, L., Greenson, J., Lord, C., Munson, J., Winter, J., and Young, G. (2014). The impact of parent-delivered intervention on parents of very young children with autism. *Journal of autism and developmental disorders*, 44(2):353–365. 13

[Fabbri-Destro et al., 2009] Fabbri-Destro, M., Cattaneo, L., Boria, S., and Rizzolatti, G. (2009). Planning actions in autism. *Experimental brain research*, 192(3):521–525. 78

[Faria et al., 2014] Faria, D. R., Premebida, C., and Nunes, U. (2014). A probabilistic approach for human everyday activities recognition using body motion from RGB-D images. *IEEE RO-MAN 2014 - 23rd IEEE International Symposium on Robot and Human Interactive Communication: Human-Robot Co-Existence: Adaptive Interfaces and Systems for Daily Life, Therapy, Assistance and Socially Engaging Interactions*, pages 732–737. 65

[Faria et al., 2015] Faria, D. R., Vieira, M., Premebida, C., and Nunes, U. (2015). Probabilistic human daily activity recognition towards robot-assisted living. In *2015 24th IEEE international symposium on robot and human interactive communication (RO-MAN)*, pages 582–587. IEEE. 65

[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*, 14(3):237–252. 15, 31

[Feil-Seifer and Mataric, 2010] Feil-Seifer, D. and Mataric, M. (2010). Using proxemics to evaluate human-robot interaction. *5th ACM/IEEE International Conference on Human-Robot Interaction, HRI 2010*. 23

[Feiner et al., 1993] Feiner, S., MacIntyre, B., and Seligmann, D. (1993). Knowledge-based augmented reality. *Communications of the ACM*, 36(7):53–62. 15

[Feldman, 1995] Feldman, L. A. (1995). Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of personality and social psychology*, 69(1):153. 98

[Finnigan and Starr, 2010] Finnigan, E. and Starr, E. (2010). Increasing social responsiveness in a child with autism: A comparison of music and non-music interventions. *Autism*, 14(4):321–348. 105

[Flanagan et al., 2012] Flanagan, J. E., Landa, R., Bhat, A., and Bauman, M. (2012). Head lag in infants at risk for autism: a preliminary study. *American Journal of Occupational Therapy*, 66(5):577–585. 23

[Forti et al., 2020] Forti, S., Colombo, B., Clark, J., Bonfanti, A., Molteni, S., Crippa, A., Antonietti, A., and Molteni, M. (2020). Soundbeam imitation intervention: Training children with autism to imitate meaningless body gestures through music. *Advances in Autism*. 106

[Fournier et al., 2010] Fournier, K. A., Hass, C. J., Naik, S. K., Lodha, N., and Cauraugh, J. H. (2010). Motor coordination in autism spectrum disorders: a synthesis and meta-analysis. *Journal of autism and developmental disorders*, 40(10):1227–1240. 17

[Fowles et al., 1981] Fowles, D. C., Christie, M. J., Edelberg, R., Grings, W. W., Lykken, D. T., and Venables, P. H. (1981). Publication recommendations for electrodermal measurements. *Psychophysiology*, 18(3):232–239. 22, 63

[Friard and Gamba, 2016] Friard, O. and Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, 7(11):1325–1330. 56

[Fu et al., 2020] Fu, Y., Zhao, J., Dong, Y., and Wang, X. (2020). Dry electrodes for human bioelectrical signal monitoring. *Sensors (Switzerland)*, 20(13):1–30. 39

[Gallo and Annenberg, 2009] Gallo, K. and Annenberg, J. (2009). Eye power 2. Thought Bubble Productions. 53

[Gepner et al., 2001] Gepner, B., Deruelle, C., and Grynfeltt, S. (2001). Motion and emotion: A novel approach to the study of face processing by young autistic children. *Journal of autism and developmental disorders*, 31(1):37–45. 17

[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. 17

[Gessaroli et al., 2013] Gessaroli, E., Santelli, E., di Pellegrino, G., and Frassinetti, F. (2013). Personal space regulation in childhood autism spectrum disorders. *PloS one*, 8(9):e74959. 100

[Gima et al., 2018] Gima, H., Kihara, H., Watanabe, H., Nakano, H., Nakano, J., Konishi, Y., Nakamura, T., and Taga, G. (2018). Early motor signs of autism spectrum disorder in spontaneous position and movement of the head. *Experimental brain research*, 236(4):1139–1148. 23

[Goffman, 1955] Goffman, E. (1955). On face-work: An analysis of ritual elements in social interaction. *Psychiatry*, 18(3):213–231. 12

[Gold et al., 2006] Gold, C., Wigram, T., and Elefant, C. (2006). Music therapy for autistic spectrum disorder. *Cochrane Database of Systematic Reviews*, (2). 107, 112

[Gordon and Fleisher, 2010] Gordon, N. J. and Fleisher, W. L. (2010). *Effective interviewing and interrogation techniques*. Academic Press. 51

[Gray and Garand, 1993] Gray, C. A. and Garand, J. D. (1993). Social stories: Improving responses of students with autism with accurate social information. *Focus on autistic behavior*, 8(1):1–10. 15

[Green et al., 2010] Green, J., Charman, T., McConachie, H., Aldred, C., Slonims, V., Howlin, P., Le Couteur, A., Leadbitter, K., Hudry, K., and Byford, S. (2010). Parent-mediated communication-focused treatment in children with autism (PACT): a randomised controlled trial. *The Lancet*, 375(9732):2152–2160. 13

[Grindle et al., 2012] Grindle, C. F., Hastings, R. P., Saville, M., Hughes, J. C., Huxley, K., Kovshoff, H., Griffith, G. M., Walker-Jones, E., Devonshire, K., and Remington, B. (2012). Outcomes of a behavioral education model for children with autism in a mainstream school setting. *Behavior modification*, 36(3):298–319. 13

[Gunes and Piccardi, 2008] Gunes, H. and Piccardi, M. (2008). Automatic temporal segment detection and affect recognition from face and body display. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 39(1):64–84. 80

[Gunes and Schüller, 2017] Gunes, H. and Schüller, B. (2017). 16 Automatic Analysis of Social Emotions. *Social Signal Processing*, page 213. 22

[Gupta et al., 2014] Gupta, S. S., Henninger, W. R., and Vinh, M. E. (2014). *First Steps to Preschool Inclusion*. Brookes Publishing, Baltimore. 97, 119

[Guy et al., 2014] Guy, L., Souders, M., Bradstreet, L., DeLussey, C., and Herrington, J. D. (2014). Brief report: Emotion regulation and respiratory sinus arrhythmia in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 44(10):2614–2620. 22

[Haas et al., 2016] Haas, K., Costley, D., Falkmer, M., Richdale, A., Sofronoff, K., and Falkmer, T. (2016). Factors influencing the research participation of adults with autism spectrum disorders. *Journal of autism and developmental disorders*, 46(5):1793–1805. 26

[Hall, 1963] Hall, E. T. (1963). A System for the notation of proxemic behavior 1. *American anthropologist*, 65(5):1003–1026. 65

[Hall et al., 2009] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*, 11(1):10–18. 72, 74, 80, 93

[Harrigan, 2013] Harrigan, J. A. (2013). Methodology: Coding and studying nonverbal behavior. 81

[Hauck et al., 1995] Hauck, M., Fein, D., Waterhouse, L., and Feinstein, C. (1995). Social initiations by autistic children to adults and other children. *Journal of autism and developmental disorders*, 25(6):579–595. 19, 61

[Heaton et al., 1999] Heaton, P., Hermelin, B., and Pring, L. (1999). Can children with autistic spectrum disorders perceive affect in music? An experimental investigation. *Psychological medicine*, 29(6):1405–1410. 105

[Hernandez et al., 2014] Hernandez, J., Riobo, I., Rozga, A., Abowd, G. D., and Picard, R. W. (2014). Using Electrodermal Activity to Recognize Ease of Engagement in Children during Social Interactions. pages 307–317. 99, 101

[Hirstein et al., 2001] Hirstein, W., Iversen, P., and Ramachandran, V. S. (2001). Autonomic responses of autistic children to people and objects. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1479):1883–1888. 4, 21

[Hopkins et al., 2011] Hopkins, I. M., Gower, M. W., Perez, T. A., Smith, D. S., Amthor, F. R., Wimsatt, F. C., and Biasini, F. J. (2011). Avatar assistant: improving social skills in students with an ASD through a computer-based intervention. *Journal of autism and developmental disorders*, 41(11):1543–1555. 15

[Hourcade et al., 2013] Hourcade, J. P., Williams, S. R., Miller, E. A., Huebner, K. E., and Liang, L. J. (2013). Evaluation of tablet apps to encourage social interaction in children with autism spectrum disorders. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 3197–3206. 14, 25, 47

[Howison et al., 2011] Howison, M., Trninic, D., Reinholz, D., and Abrahamson, D. (2011). The Mathematical Imagery Trainer: from embodied interaction to conceptual learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1989–1998. 2, 17

[Hubert et al., 2009] Hubert, B. E., Wicker, B., Monfardini, E., and Deruelle, C. (2009). Electrodermal reactivity to emotion processing in adults with autistic spectrum disorders. *Autism*, 13(1):9–19. 22

[Illemsen et al., 2009] Illemsen, A. L. W., Eyer, E. R. I. K. S. M., and Oldus, L. U. P. J. J. N. (2009). The Observer XT : A tool for the integration and synchronization of multimodal signals. 41(3):731–735. 56

[Ingersoll, 2008] Ingersoll, B. (2008). The Social Role of Imitation in Autism: Implications for the Treatment of Imitation Deficits. *Infants & Young Children*, 21(2):107–119. 14, 86

[Jaimes and Sebe, 2007] Jaimes, A. and Sebe, N. (2007). Multimodal human–computer interaction: A survey. *Computer vision and image understanding*, 108(1-2):116–134. 2, 25

[James et al., 2015] James, R., Sigafoos, J., Green, V. A., Lancioni, G. E., O'Reilly, M. F., Lang, R., Davis, T., Carnett, A., Achmadi, D., and Gevarter, C. (2015). Music therapy for individuals with autism spectrum disorder: A systematic review. *Review Journal of Autism and Developmental Disorders*, 2(1):39–54. 105

[Jaschke, 2014] Jaschke, A. C. (2014). Music intervention as system: reversing hyper systemising in autism spectrum disorders to the comprehension of music as intervention. *Medical hypotheses*, 82(1):40–48. 108

[Javed et al., 2020] Javed, H., Lee, W., and Park, C. H. (2020). Toward an Automated Measure of Social Engagement for Children With Autism Spectrum Disorder—A Personalized Computational Modeling Approach. *Frontiers in Robotics and AI*, 7:43. 26

[Johnston et al., 2006] Johnston, J. M., Foxx, R. M., Jacobson, J. W., Green, G., and Mulick, J. A. (2006). Positive behavior support and applied behavior analysis. *The Behavior Analyst*, 29(1):51–74. 13

[Joseph et al., 2008] Joseph, R. M., Ehrman, K., McNally, R., and Keehn, B. (2008). Affective response to eye contact and face recognition ability in children with ASD. In *Journal of the International Neuropsychological Society*. 22

[Kasari et al., 2008] Kasari, C., Paparella, T., Freeman, S., and Jahromi, L. B. (2008). Language outcome in autism: randomized comparison of joint attention and play interventions. *Journal of consulting and clinical psychology*, 76(1):125. 13

[Kasari et al., 2012] Kasari, C., Rotheram-Fuller, E., Locke, J., and Gulsrud, A. (2012). Making the connection: Randomized controlled trial of social skills at school for children with autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, 53(4):431–439. 35

[Kasos et al., 2020] Kasos, K., Kekecs, Z., Csirmaz, L., Zimonyi, S., Vikor, F., Kasos, E., Veres, A., Kotyuk, E., and Szekely, A. (2020). Bilateral comparison of traditional and alternate electrodermal measurement sites. *Psychophysiology*, 57(11):e13645. 41

[Ke and Im, 2013] Ke, F. and Im, T. (2013). Virtual-reality-based social interaction training for children with high-functioning autism. *The Journal of Educational Research*, 106(6):441–461. 15

[Keay-Bright, 2007] Keay-Bright, W. (2007). The Reactive Colours Project: Demonstrating Participatory and Collaborative Design Methods for the Creation of Software for Autistic Children. 17

[Kerr et al., 2002] Kerr, S. J., Neale, H. R., and Cobb, S. V. G. (2002). Virtual environments for social skills training: the importance of scaffolding in practice. In *Proceedings of the fifth international ACM conference on Assistive technologies*, pages 104–110. 15

[Kirby et al., 2015] Kirby, A. V., Dickie, V. A., and Baranek, G. T. (2015). Sensory experiences of children with autism spectrum disorder: In their own words. *Autism*, 19(3):316–326. 24, 25, 47

[Kleinsmith et al., 2005] Kleinsmith, A., De Silva, P. R., and Bianchi-Berthouze, N. (2005). Grounding affective dimensions into posture features. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3784 LNCS:263–270. 65

[Klin et al., 2005] Klin, A., Jones, W., Schultz, R. T., and Volkmar, F. R. (2005). The Enactive Mind-From Actions to Cognition: Lessons from Autism. 12

[Klusek et al., 2015] Klusek, J., Roberts, J. E., and Losh, M. (2015). Cardiac autonomic regulation in autism and Fragile X syndrome: a review. *Psychological Bulletin*, 141(1):141. 22

[Knott et al., 1995] Knott, F., Lewis, C., and Williams, T. (1995). Sibling interaction of children with learning disabilities: A comparison of autism and Down's syndrome. *Journal of child psychology and psychiatry*, 36(6):965–976. 86

[Knott et al., 2007] Knott, F., Lewis, C., and Williams, T. (2007). Sibling interaction of children with autism: Development over 12 months. *Journal of Autism and Developmental Disorders*, 37(10):1987–1995. 86

[Koo et al., 2020] Koo, S. H., Gaul, K., Rivera, S., Pan, T., and Fong, D. (2020). Wearable Technology Design for Autism Spectrum Disorders. (February 2018). 37, 118

[Kushki et al., 2013] Kushki, A., Drumm, E., Mobarak, M. P., Tanel, N., Dupuis, A., Chau, T., and Anagnostou, E. (2013). Investigating the autonomic nervous system response to anxiety in children with autism spectrum disorders. *PLoS one*, 8(4):e59730. 4, 21

[Kylliäinen and Hietanen, 2006] Kylliäinen, A. and Hietanen, J. K. (2006). Skin conductance responses to another person's gaze in children with autism. *Journal of autism and developmental disorders*, 36(4):517–525. 22

[Laborde et al., 2017] Laborde, S., Mosley, E., and Thayer, J. F. (2017). Heart rate variability and cardiac vagal tone in psychophysiological research - Recommendations for experiment planning, data analysis, and data reporting. *Frontiers in Psychology*, 8(FEB):1–18. 22, 64, 98

[Lang, 1995] Lang, P. J. (1995). The emotion probe: studies of motivation and attention. *American psychologist*, 50(5):372. 22

[Lanni et al., 2012] Lanni, K. E., Schupp, C. W., Simon, D., and Corbett, B. A. (2012). Verbal ability, social stress, and anxiety in children with autistic disorder. *Autism*, 16(2):123–138. 47

[LeGoff, 2004] LeGoff, D. B. (2004). Use of LEGO© as a therapeutic medium for improving social competence. *Journal of autism and developmental disorders*, 34(5):557–571. 13, 37

[Legoff and Sherman, 2006] Legoff, D. B. and Sherman, M. (2006). Long-term outcome of social skills intervention based on interactive LEGO© play. *Autism*, 10(4):317–329. 13

[Leman, 2008] Leman, M. (2008). *Embodied music cognition and mediation technology*. MIT press. 106

[Levine et al., 2012] Levine, T. P., Sheinkopf, S. J., Pescosolido, M., Rodino, A., Elia, G., and Lester, B. (2012). Physiologic arousal to social stress in children with autism spectrum disorders: A pilot study. *Research in Autism Spectrum Disorders*, 6(1):177–183. 22

[Licht et al., 2009] Licht, C. M. M., De Geus, E. J. C., Van Dyck, R., and Penninx, B. W. J. H. (2009). Association between anxiety disorders and heart rate variability in The Netherlands Study of Depression and Anxiety (NESDA). *Psychosomatic medicine*, 71(5):508–518. 22

[Liebowitz, 1987] Liebowitz, M. R. (1987). Social Phobia. In *Modern Problems of Pharmacopsychiatry*, pages 141–173. 57

[Liu et al., 2008] Liu, C., Conn, K., Sarkar, N., and Stone, W. (2008). Physiology-based affect recognition for computer-assisted intervention of children with Autism Spectrum Disorder. *International journal of human-computer studies*, 66(9):662–677. 25

[Liu et al., 2017] Liu, W., Zhou, T., Zhang, C., Zou, X., and Li, M. (2017). Response to name: A dataset and a multimodal machine learning framework towards autism study. In *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 178–183. IEEE. 26

[Lord et al., 1999] Lord, C., Rutter, M., and DiLavore, P. C. (1999). Autism Diagnostic Observation Schedule–Generic. *Dissertation Abstracts International Section A: Humanities and Social Sciences*. 20, 50

[Mahoney and Perales, 2003] Mahoney, G. and Perales, F. (2003). Using relationship-focused intervention to enhance the social—emotional functioning of young children with autism spectrum disorders. *Topics in Early Childhood Special Education*, 23(2):74–86. 13

[Malik et al., 1996] Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J., and Schwartz, P. J. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European heart journal*, 17(3):354–381. 74

[Malik and Lindahl, 2004] Malik, N. M. and Lindahl, K. M. (2004). System for coding interactions in dyads (SCID). *Couple observational coding systems*, pages 173–188. 18

[Malinverni et al., 2014] Malinverni, L., Mora-Guiard, J., Padillo, V., Mairena, M., Hervás, A., and Pares, N. (2014). Participatory design strategies to enhance the creative contribution of children with special needs. In *Proceedings of the 2014 conference on Interaction design and children - IDC '14*, pages 85–94, New York, New York, USA. ACM Press. 18

[Marinoiu et al., 2018] Marinoiu, E., Zanfir, M., Olaru, V., and Sminchisescu, C. (2018). 3d human sensing, action and emotion recognition in robot assisted therapy of children with autism. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2158–2167. 26

[Martin et al., 2018] Martin, K. B., Hammal, Z., Ren, G., Cohn, J. F., Cassell, J., Ogihara, M., Britton, J. C., Gutierrez, A., and Messinger, D. S. (2018). Objective measurement of head movement differences in children with and without autism spectrum disorder. *Molecular autism*, 9(1):1–10. 23

[Marwecki et al., 2013] Marwecki, S., Rädle, R., and Reiterer, H. (2013). Encouraging collaboration in hybrid therapy games for autistic children. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13*, page 469, New York, New York, USA. ACM Press. 15, 31

[Mason et al., 2005] Mason, M. F., Tatkow, E. P., and Macrae, C. N. (2005). The look of love: Gaze shifts and person perception. *Psychological science*, 16(3):236–239. 65

[Mathersul et al., 2013] Mathersul, D., McDonald, S., and Rushby, J. A. (2013). Automatic facial responses to briefly presented emotional stimuli in autism spectrum disorder. *Biological Psychology*, 94(2):397–407. 22

[Mendes, 2009] Mendes, W. B. (2009). Assessing autonomic nervous system activity. *Methods in social neuroscience*, 118:147. 21

[Messinger et al., 2015] Messinger, D. S., Duvivier, L. L., Warren, Z. E., Mahoor, M., Baker, J., Warlaumont, A., and Ruvolo, P. (2015). Affective computing, emotional development, and autism. 25

[Migueles et al., 2017] Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Nyström, C. D., Mora-Gonzalez, J., Löf, M., Labayen, I., Ruiz, J. R., and Ortega, F. B. (2017). Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. *Sports medicine*, 47(9):1821–1845. 64

[Millen et al., 2010] Millen, L., Edlin-White, R., and Cobb, S. (2010). The development of educational collaborative virtual environments for children with autism. In *Proceedings of the 5th Cambridge Workshop on Universal Access and Assistive Technology, Cambridge*, volume 1, page 7. 112

[Mitchell et al., 2007] Mitchell, P., Parsons, S., and Leonard, A. (2007). Using virtual environments for teaching social understanding to 6 adolescents with autistic spectrum disorders. *Journal of autism and developmental disorders*, 37(3):589–600. 15

[Moore and Calvert, 2000] Moore, M. and Calvert, S. (2000). Brief Report: Vocabulary Acquisition for Children with Autism: Teacher or Computer Instruction. *Journal of Autism and Developmental Disorders*, 30(4):359–362. 2, 14

[Mora Guiard, 2009] Mora Guiard, J. (2009). SIIMTA, Sistema Integrat d'Interacció de cos sencer generada en temps real i MusicoTeràpia per a persones amb discapacitats com l'Autisme. 17, 112

[Mora-Guiard et al., 2016] Mora-Guiard, J., Crowell, C., Pares, N., and Heaton, P. (2016). Lands of Fog: helping children with Autism in social interaction through a full-body interactive experience. *ACM SIGCHI Conference on Interaction Design and Children 2016*, page manuscript submitted for publication. 18, 32

[Morawska et al., 2015] Morawska, A., Basha, A., Adamson, M., and Winter, L. (2015). Microanalytic coding versus global rating of maternal parenting behaviour. *Early Child Development and Care*, 185(3):448–463. 19

[Myles et al., 2001] Myles, B. S., Barnhill, G. P., Hagiwara, T., Griswold, D. E., and Simpson, R. L. (2001). A synthesis of studies on the intellectual, academic, social/emotional and sensory characteristics of children and youth with Asperger syndrome. *Education and Training in Mental Retardation and Developmental Disabilities*, pages 304–311. 13

[Neuhaus et al., 2014] Neuhaus, E., Bernier, R., and Beauchaine, T. P. (2014). Brief report: Social skills, internalizing and externalizing symptoms, and respiratory sinus arrhythmia in autism. *Journal of Autism and Developmental Disorders*, 44(3):730–737. 22

[Neuhaus et al., 2016] Neuhaus, E., Bernier, R. A., and Beauchaine, T. P. (2016). Children with autism show altered autonomic adaptation to novel and familiar social partners. *Autism Research*, 9(5):579–591. 21, 72

[Neumann and Blanton, 1970] Neumann, E. and Blanton, R. (1970). The early history of electrodermal research. *Psychophysiology*, 6(4):453–475. 38

[Nikopoulos and Keenan, 2003] Nikopoulos, C. K. and Keenan, M. (2003). Promoting social initiation in children with autism using video modeling. *Behavioral Interventions*, 18(2):87–108. 12

[Ortega et al., 2015] Ortega, D. H., Cibrian, F. L., and Tentori, M. (2015). BendableSound: a fabric-based interactive surface to promote free play in children with autism. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*, pages 315–316. 15, 31, 106

[Osr, 2016] Osr (2016). OpenSignals (r)evolution Software Data Sheet. Technical report. xv, 44, 45

[Otzenberger et al., 1998] Otzenberger, H., Gronfier, C., Simon, C., Charloux, a., Ehrhart, J., Piquard, F., and Brandenberger, G. (1998). Dynamic heart rate variability: a tool for exploring sympathovagal balance continuously during sleep in men. *The American journal of physiology*, 275(3 Pt 2):H946–H950. 74

[Overy, 2003] Overy, K. (2003). Dyslexia and music: From timing deficits to musical intervention. *Annals of the New York academy of sciences*, 999(1):497–505. 105

[Overy and Molnar-Szakacs, 2009] Overy, K. and Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music perception*, 26(5):489–504. 106

[Ozonoff et al., 2010] Ozonoff, S., Iosif, A.-M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., Rogers, S. J., Rozga, A., Sangha, S., and Sigman, M. (2010). A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(3):256–266. 23

[Pajareya and Nopmaneejumruslers, 2011] Pajareya, K. and Nopmaneejumruslers, K. (2011). A pilot randomized controlled trial of DIR/Floortime™ parent training intervention for pre-school children with autistic spectrum disorders. *Autism : the international journal of research and practice*, 15(5):563–77. 13

[Papert, 1980] Papert, S. (1980). Mindstorms: children, computers, and powerful ideas Basic Books. *Inc. New York, NY*. 34, 100

[Parés and Altimira, 2013] Parés, N. and Altimira, D. (2013). Analyzing the adequacy of interaction paradigms in artificial reality experiences. *Human-Computer Interaction*, 28(2):77–114. 18

[Pares et al., 2005] Pares, N., Masri, P., Van Wolferen, G., and Creed, C. (2005). Achieving dialogue with children with severe autism in an adaptive multisensory interaction: the" MEDIATE" project. *IEEE Transactions on Visualization and Computer Graphics*, 11(6):734–743. 2, 17

[Parés and Parés, 2001] Parés, N. and Parés, R. (2001). Interaction-driven virtual reality application design: A particular case: El Ball del Fanalet or Lightpools. *Presence: Teleoperators and Virtual Environments*, 10(2):236–245. 30

[Parsons and Mitchell, 2002] Parsons, S. and Mitchell, P. (2002). The potential of virtual reality in social skills training for people with autistic spectrum disorders. *Journal of intellectual disability research*, 46(5):430–443. 15

[Parten, 1932] Parten, M. B. (1932). Social participation among pre-school children. *The Journal of Abnormal and Social Psychology*, 27(3):243–269. 19

[Patriquin et al., 2013] Patriquin, M. A., Scarpa, A., Friedman, B. H., and Porges, S. W. (2013). Respiratory sinus arrhythmia: A marker for positive social functioning and receptive language skills in children with autism spectrum disorders. *Developmental psychobiology*, 55(2):101–112. 22

[Peterson et al., 2015] Peterson, C. C., Slaughter, V., and Brownell, C. (2015). Children with autism spectrum disorder are skilled at reading emotion body language. *Journal of experimental child psychology*, 139:35–50. 78

[Pfeiffer et al., 2005] Pfeiffer, B., Kinnealey, M., Reed, C., and Herzberg, G. (2005). Sensory modulation and affective disorders in children and adolescents with Asperger's disorder. *American Journal of Occupational Therapy*, 59(3):335–345. 13

[Piaget, 2013] Piaget, J. (2013). *Play, dreams and imitation in childhood*. 19

[Piana et al., 2013] Piana, S., Staglianò, A., Camurri, A., and Odone, F. (2013). A set of full-body movement features for emotion recognition to help children affected by autism spectrum condition. In *IDGEI International Workshop*. 23

[Picard, 2009] Picard, R. W. (2009). Future affective technology for autism and emotion communication. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1535):3575–84. 21, 37, 38, 118

[Picard and Healey, 1997] Picard, R. W. and Healey, J. (1997). Affective wearables. *Personal Technologies*, 1(4):231–240. 21

[Ploog et al., 2013] Ploog, B. O., Scharf, A., Nelson, D., and Brooks, P. J. (2013). Use of computer-assisted technologies (CAT) to enhance social, communicative, and language development in children with autism spectrum disorders. *Journal of autism and developmental disorders*, 43(2):301–22. 12

[PLUX Wireless Biosignals S.A.,] PLUX Wireless Biosignals S.A. biosignalsplux Explorer - PLUX Store. 39

[Porayska-Pomsta et al., 2012] Porayska-Pomsta, K., Frauenberger, C., Pain, H., Rajendran, G., Smith, T., Menzies, R., Foster, M. E., Alcorn, A., Wass, S., and Bernadini, S. (2012). Developing technology for autism: an interdisciplinary approach. *Personal and Ubiquitous Computing*, 16(2):117–127. 2, 17

[Porges, 2007] Porges, S. W. (2007). The polyvagal perspective. *Biological psychology*, 74(2):116–143. 22

[Porges et al., 1994] Porges, S. W., Doussard-Roosevelt, J. A., and Maiti, A. K. (1994). Vagal tone and the physiological regulation of emotion. *Monographs of the society for research in child development*, 59(2-3):167–186. 22

[Posada-Quintero and Chon, 2020] Posada-Quintero, H. F. and Chon, K. H. (2020). Innovations in electrodermal activity data collection and signal processing: A systematic review. *Sensors (Switzerland)*, 20(2). 41, 43, 64, 99

[Provost et al., 2007] Provost, B., Lopez, B. R., and Heimerl, S. (2007). A comparison of motor delays in young children: autism spectrum disorder, developmental delay, and developmental concerns. *Journal of autism and developmental disorders*, 37(2):321–328. 78

[Ragone, 2020] Ragone, G. (2020). Designing Embodied Musical Interaction for Children with Autism. *ASSETS 2020 - 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. 106

[Ramdoss et al., 2012] Ramdoss, S., MacHalicek, W., Rispoli, M., Mulloy, A., Lang, R., and O'Reilly, M. (2012). Computer-based interventions to improve social and emotional skills in individuals with autism spectrum disorders: A systematic review. *Developmental Neurorehabilitation*, 15(2):119–135. 14

[Rao et al., 2008] Rao, P. A., Beidel, D. C., and Murray, M. J. (2008). Social skills interventions for children with Asperger's syndrome or high-functioning autism: a review and recommendations. *Journal of autism and developmental disorders*, 38(2):353–61. 1, 13

[Read et al., 2002] Read, J., Macfarlane, S., and Casey, C. (2002). Endurability , Engagement and Expectations : Measuring Children ' s Fun. *Interaction Design and Children*, 2(January):1–23. 24, 48

[Rehg et al., 2013] Rehg, J., Abowd, G., Rozga, A., Romero, M., Clements, M., Sclaroff, S., Essa, I., Ousley, O., Li, Y., and Kim, C. (2013). Decoding children's social behavior. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3414–3421. 1, 23

[Rehg, 2011] Rehg, J. M. (2011). Behavior Imaging: Using Computer Vision to Study Autism. *MVA*, 11:14–21. 23

[Rehg et al., 2014] Rehg, J. M., Rozga, A., Abowd, G. D., and Goodwin, M. S. (2014). Behavioral Imaging and Autism. pages 84–87. 26

[Riby et al., 2012] Riby, D. M., Whittle, L., and Doherty-Sneddon, G. (2012). Physiological reactivity to faces via live and video-mediated communication in typical and atypical development. *Journal of Clinical and Experimental Neuropsychology*, 34(4):385–395. 22

[Rinehart et al., 2006] Rinehart, N. J., Tonge, B. J., Iansek, R., McGinley, J., Brereton, A. V., Enticott, P. G., and Bradshaw, J. L. (2006). Gait function in newly diagnosed children with autism: cerebellar and basal ganglia related motor disorder. *Developmental medicine and child neurology*, 48(10):819–824. 78

[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. 14

[Rogers et al., 2003] Rogers, S. J., Hepburn, S. L., Stackhouse, T., and Wehner, E. (2003). Imitation performance in toddlers with autism and those with other developmental disorders. *Journal of child psychology and psychiatry*, 44(5):763–781. 86

[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. 17

[Rottenberg, 2007] Rottenberg, J. (2007). Cardiac vagal control in depression: a critical analysis. *Biological psychology*, 74(2):200–211. 22

[Roussos et al., 1999] Roussos, M., Johnson, A., Moher, T., Leigh, J., Vasilakis, C., and Barnes, C. (1999). Learning and Building Together in an Immersive Virtual World. *Presence: Teleoperators and Virtual Environments*, 8(3):247–263. 16

[Rowland and Schweigert, 2009] Rowland, C. M. and Schweigert, P. D. (2009). Object lessons: How children with autism spectrum disorders use objects to interact with the physical and social environments. *Research in Autism Spectrum Disorders*, 3(2):517–527. 13

[Rubin, 2008] Rubin, K. H. (2008). THE PLAY OBSERVAT ON SCALE POS I (). 20

[Rudovic et al., 2018] Rudovic, O., Lee, J., Dai, M., Schuller, B., and Picard, R. W. (2018). Personalized machine learning for robot perception of affect and engagement in autism therapy. *Science Robotics*, 3(19). 23, 25

[Samuels, 2014] Samuels, K. (2014). Enabling creativity: Inclusive music interfaces and practices. In *Proceedings of International Conference on Live Interfaces (ICLI)*. 106

[Sanghvi et al., 2011] Sanghvi, J., Castellano, G., Leite, I., Pereira, A., McOwan, P. W., and Paiva, A. (2011). Automatic analysis of affective postures and body motion to detect engagement with a game companion. In *Proceedings of the 6th international conference on Human-robot interaction*, pages 305–312. ACM. 23

[Sayis et al., 2019] Sayis, B., Crowell, C., Benitez, J., Ramirez, R., and Pares, N. (2019). Computational Modeling of Psycho-physiological Arousal and Social Initiation of children with Autism in Interventions through Full-Body Interaction. In *2019 8th International Conference on Affective Computing and Intelligent Interaction, ACII 2019*. 117

[Sayis et al., 2020] Sayis, B., Pares, N., and Gunes, H. (2020). Bodily Expression of Social Initiation Behaviors in ASC and non-ASC children: Mixed Reality vs. LEGO Game Play. 117

[Scassellati, 2005] Scassellati, B. (2005). Quantitative metrics of social response for autism diagnosis. In *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005.*, pages 585–590. IEEE. 14, 26

[Scassellati et al., 2012] Scassellati, B., Admoni, H., and Matarić, M. (2012). Robots for use in autism research. *Annual review of biomedical engineering*, 14:275–294. 14

[Schuller et al., 2013] Schuller, B., Marchi, E., Baron-Cohen, S., O'Reilly, H., Robinson, P., Davies, I., Golan, O., Friedenson, S., Tal, S., and Newman, S. (2013). ASC-Inclusion: Interactive emotion games for social inclusion of children with Autism Spectrum Conditions. In *Proceedings 1st International Workshop on Intelligent Digital Games for Empowerment and Inclusion (IDGEI 2013) held in conjunction with the 8th Foundations of Digital Games 2013 (FDG)(B. Schuller, L. Paletta, and N. Sabouret, eds.), Chania, Greece*. 15, 23

[Sensetoys, 2021] Sensetoys (2021). sensory weighted shoulder weight. xv, 41

[Sequeira et al., 2009] Sequeira, H., Hot, P., Silvert, L., and Delplanque, S. (2009). Electrical autonomic correlates of emotion. *International journal of psychophysiology*, 71(1):50–56. 22

[Sheikhi and Odobez, 2012] Sheikhi, S. and Odobez, J.-M. (2012). Recognizing the visual focus of attention for human robot interaction. In *International Workshop on Human Behavior Understanding*, pages 99–112. Springer. 65

[Sigman et al., 1999] Sigman, M., Ruskin, E., Arbelle, S., Corona, R., Dissanayake, C., Espinosa, M., Kim, N., López, A., Zierhut, C., and Mervis, C. B. (1999). Continuity and change in the social competence of children with autism, Down syndrome, and developmental delays. *Monographs of the society for research in child development*, pages i–139. 12, 72, 93

[Sigrist et al., 2015] Sigrist, R., Rauter, G., Marchal-Crespo, L., Riener, R., and Wolf, P. (2015). Sonification and haptic feedback in addition to visual feedback enhances complex motor task learning. *Experimental brain research*, 233(3):909–925. 106

[Sluis et al., 2008] Sluis, F. V. D., Dijk, E. M. a. G. V., and Perloy, L. M. (2008). Measuring Fun and Enjoyment of Children in a Museum : Evaluating the Smileyometer Study One : Prototype. *Proceding of Measuring*, 2012:86–89. 24

[South et al., 2011] South, M., Dana, J., White, S. E., and Crowley, M. J. (2011). Failure is not an option: Risk-taking is moderated by anxiety and also by cognitive ability in children and adolescents diagnosed with an autism spectrum disorder. *Journal of autism and developmental disorders*, 41(1):55–65. 22

[Spielberger, 2010] Spielberger, C. D. (2010). State-Trait Anxiety Inventory. In *The Corsini Encyclopedia of Psychology*, pages 1–1. John Wiley & Sons, Inc., Hoboken, NJ, USA. 4, 47

[Spss, 2011] Spss, I. (2011). IBM SPSS statistics for Windows, version 20.0. *New York: IBM Corp*, 440. 89

[Srinivasan et al., 2016] Srinivasan, S. M., Eigsti, I. M., Gifford, T., and Bhat, A. N. (2016). The effects of embodied rhythm and robotic interventions on the spontaneous and responsive verbal communication skills of children with Autism Spectrum Disorder (ASD): A further outcome of a pilot randomized controlled trial. *Research in Autism Spectrum Disorders*, 27:73–87. 2, 14, 106

[Stahl, 2013] Stahl, B. C. (2013). Responsible research and innovation: The role of privacy in an emerging framework. *Science and Public Policy*, 40(6):708–716. 11

[Stone et al., 1997] Stone, W. L., Ousley, O. Y., and Littleford, C. D. (1997). Motor imitation in young children with autism: What's the object? *Journal of abnormal child psychology*, 25(6):475–485. 86

[Strain, 1983] Strain, P. S. (1983). Generalization of autistic children's social behavior change: Effects of developmentally integrated and segregated settings. *Analysis and Intervention in Developmental Disabilities*, 3(1):23–34. 97, 119

[Strain et al., 1979] Strain, P. S., Kerr, M. M., and Ragland, E. U. (1979). Effects of peer-mediated social initiations and prompting/reinforcement procedures on the social behavior of autistic children. *Journal of autism and developmental disorders*, 9(1):41–54. 2, 13

[Streeck et al., 2011] Streeck, J., Goodwin, C., and LeBaron, C. (2011). *Embodied interaction: Language and body in the material world*. Cambridge University Press. 25

[Sturgess et al., 2002] Sturgess, J., Rodger, S., and Ozanne, A. (2002). A review of the use of self-report assessment with young children. *British Journal of Occupational Therapy*, 65(3):108–116. 24

[Sturgess and Ziviani, 1996] Sturgess, J. and Ziviani, J. (1996). A self-report play skills questionnaire: Technical development. *Australian Occupational Therapy Journal*, 43(3-4):142–154. 24

[St-Laurent-Gagnon et al., 1999] St-Laurent-Gagnon, T., Bernard-Bonnin, A. C., and Villeneuve, E. (1999). Pain evaluation in preschool children and by their parents. *Acta Paediatrica*, 88(4):422–427. 24

[Symes and Humphrey, 2011] Symes, W. and Humphrey, N. (2011). School factors that facilitate or hinder the ability of teaching assistants to effectively support pupils with autism spectrum disorders (ASDs) in mainstream secondary schools. *Journal of Research in Special Educational Needs*, 11(3):153–161. 35

[Taj-Eldin et al., 2018] Taj-Eldin, M., Ryan, C., O'flynn, B., and Galvin, P. (2018). A review of wearable solutions for physiological and emotional monitoring for use by people with autism spectrum disorder and their caregivers. *Sensors (Switzerland)*, 18(12). 37, 118

[Tajadura-Jiménez et al., 2017] Tajadura-Jiménez, A., Väljamäe, A., Bevilacqua, F., and Bianchi-Berthouze, N. (2017). *Principles for Designing Body-Centered Auditory Feedback*, volume 1. 16

[Tallal and Gaab, 2006] Tallal, P. and Gaab, N. (2006). Dynamic auditory processing, musical experience and language development. *Trends in neurosciences*, 29(7):382–390. 105

[Tardif et al., 1995] Tardif, C., Plumet, M.-H., Beaudichon, J., Waller, D., Bouvard, M., and Leboyer, M. (1995). Micro-analysis of social interactions between autistic children and normal adults in semi-structured play situations. *International Journal of Behavioral Development*, 18(4):727–747. 19

[Tarvainen et al., 2014] Tarvainen, M. P., Niskanen, J.-P., Lipponen, J. A., Ranta-Aho, P. O., and Karjalainen, P. A. (2014). Kubios HRV–heart rate variability analysis software. *Computer methods and programs in biomedicine*, 113(1):210–220. 64

[Taylor, S., and Jaques,] Taylor, S., and Jaques, N. EDA Explorer. 63

[Teitelbaum et al., 1998] Teitelbaum, P., Teitelbaum, O., Nye, J., Fryman, J., and Maurer, R. G. (1998). Movement analysis in infancy may be useful for early diagnosis of autism. *Proceedings of the National Academy of Sciences*, 95(23):13982–13987. 78

[Thaut and Hoemberg, 2014] Thaut, M. and Hoemberg, V. (2014). *Handbook of neurologic music therapy*. Oxford University Press (UK). 108

[Tolentino et al., 2009] Tolentino, L., Birchfield, D., and Kelliher, A. (2009). SMALLab for special needs: Using a mixed-reality platform to explore learning for children with autism. In *Proceedings of the NSF Media Arts, Science and Technology Conference, Santa Barbara, CA, USA*, pages 29–30. 2, 17

[Torres et al., 2013] Torres, E. B., Nguyen, J., Suresh, C., Yanovich, P., and Kolevzon, A. (2013). Noise from the periphery in autism spectrum disorders of idiopathic origins and of known etiology. In *Paper Presentation at the Annual Meeting of the Society for Neuroscience SFN*. 78

[Turpin et al., 1983] Turpin, G., Shine, P., and Lader, M. H. (1983). Ambulatory electrodermal monitoring: effects of ambient temperature, general activity, electrolyte media, and length of recording. *Psychophysiology*. 76

[van der Geest et al., 2002] van der Geest, J. N., Kemner, C., Camfferman, G., Verbaten, M. N., and van Engeland, H. (2002). Looking at images with human figures: comparison between autistic and normal children. *Journal of autism and developmental disorders*, 32(2):69–75. 12

[van Dooren et al., 2012] van Dooren, M., de Vries, J. J. G. G. J., and Janssen, J. H. (2012). Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology and Behavior*, 106(2):298–304. 38

[Villafuerte et al., 2012] Villafuerte, L., Markova, M., and Jorda, S. (2012). Acquisition of social abilities through musical tangible user interface: children with autism spectrum condition and the reactable. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pages 745–760. 15, 31

[Vinciarelli et al., 2009] Vinciarelli, A., Pantic, M., and Bourlard, H. (2009). Social signal processing: Survey of an emerging domain. *Image and vision computing*, 27(12):1743–1759. 1, 22

[Vygotsky, 2012] Vygotsky, L. S. (2012). *Thought and language*. MIT press. 16

[Wan et al., 2010] Wan, C. Y., Demaine, K., Zipse, L., Norton, A., and Schlaug, G. (2010). From music making to speaking: engaging the mirror neuron system in autism. *Brain research bulletin*, 82(3-4):161–168. 105

[Wasilewski and Poloński, 2012] Wasilewski, J. and Poloński, L. (2012). An introduction to ECG interpretation. In *ECG Signal Processing, Classification and Interpretation*, pages 1–20. Springer. 38

[Wechsler, 1949] Wechsler, D. (1949). Wechsler intelligence scale for children. 51

[Welch, 2012] Welch, K. C. (2012). Physiological signals of autistic children can be useful. *IEEE Instrumentation & Measurement Magazine*, 15(1):28–32. 21

[Welch et al., 2010] Welch, K. C., Lahiri, U., Warren, Z., and Sarkar, N. (2010). An approach to the design of socially acceptable robots for children with autism spectrum disorders. *International Journal of Social Robotics*, 2(4):391–403. 25

[Wertsch, 1983] Wertsch, J. V. (1983). The concept of activity in Soviet psychology. 16

[White et al., 2014] White, S. W., Mazefsky, C. A., Dichter, G. S., Chiu, P. H., Richey, J. A., and Ollendick, T. H. (2014). Social-cognitive, physiological, and neural mechanisms underlying emotion regulation impairments: Understanding anxiety in autism spectrum disorder. *International Journal of Developmental Neuroscience*, 39:22–36. 22

[White et al., 2009] White, S. W., Oswald, D., Ollendick, T., and Scahill, L. (2009). Anxiety in children and adolescents with autism spectrum disorders. *Clinical Psychology Review*, 29(3):216–229. 13

[Whyatt and Craig, 2013] Whyatt, C. and Craig, C. (2013). Sensory-motor problems in Autism. *Frontiers in integrative neuroscience*, 7:51. 78

[Whyatt and Craig, 2012] Whyatt, C. P. and Craig, C. M. (2012). Motor skills in children aged 7–10 years, diagnosed with autism spectrum disorder. *Journal of autism and developmental disorders*, 42(9):1799–1809. 17

[Williams, 2003] Williams, E. (2003). A comparative review of early forms of object-directed play and parent-infant play in typical infants and young children with autism. *Autism : the international journal of research and practice*, 7(4):361–77. 13

[Wilson, 2002] Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4):625–36. 2, 16

[Winoto et al., 2016] Winoto, P., Chen, C. G., and Tang, T. Y. (2016). The development of a Kinect-based online socio-meter for users with social and communication skill impairments: a computational sensing approach. In *2016 IEEE International Conference on Knowledge Engineering and Applications (ICKEA)*, pages 139–143. IEEE. 24

[Zachor and Ben Itzchak, 2010] Zachor, D. A. and Ben Itzchak, E. (2010). Treatment approach, autism severity and intervention outcomes in young children. *Research in Autism Spectrum Disorders*, 4(3):425–432. 1, 13

Appendix A

APPENDIX

- Figure A.1 Relaxation exercises
- Figure A.2 Visual support schedule of the experimental session
- Figure A.3 Diary of the experiment
- Figure A.4 Ethics committee approval
- Figure A.5 Multitouch tablet questionnaire items

RELAXATION EXERCISES FOR CHILDREN WITH ASC

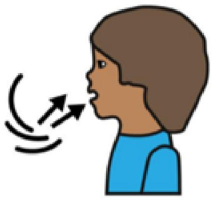
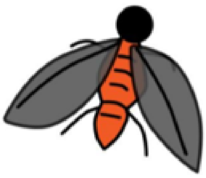
Take deep breath 	**Deep Breathing:** Take a deep breath, hold the breath for a few seconds and then release it. For young children, using a pinwheel or bubbles to practice can help them focus on their breathing and distract from their stress
Squeeze a lemon 	**Hands and Arms - Squeeze a Lemon:** Pretend you have a whole lemon in each hand. Now squeeze it hard. Try to squeeze all the juice out! Feel the tightness in your hand and arm as you squeeze. Squeeze hard! Don't leave a single drop. (Hold for 10 seconds.) Now relax and let the lemon drop from your hand. See how much better your hand and arm feel when they are relaxed.
Get that fly off your nose 	**Face and Nose - Get That Fly off Your Nose:** Here comes a pesky old fly and he has landed on your nose! Try to get him off without using your hands. Wrinkle up your nose. Make as many wrinkles in your nose as you can. Scrunch up your nose real hard and hold it just as tight as you can. Notice that when you scrunch up your nose, your cheeks and your mouth and your forehead and your eyes all help you and they get tight too. (Hold for 10 seconds.) Good. You've chased him away. Now you can just relax and let your whole face go smooth.
Squish your toes in the mud 	**Legs and Feet - Squish your Toes in the Mud:** Now pretend that you are standing barefoot in a big, fat mud puddle. Squish your toes down deep into the mud. Try to get your feet down to the bottom of the mud puddle. You'll probably need your legs to help you push. Squish your toes down. Push your feet, hard! (Hold for 10 seconds.) Okay, come back out now. Relax your feet, relax your legs, and relax your toes. It feels so good to be relaxed. No tenseness anywhere. You feel warm and tingly.

Figure A.1: Relaxation Exercises for children with ASC [Autismspeaks.org,]

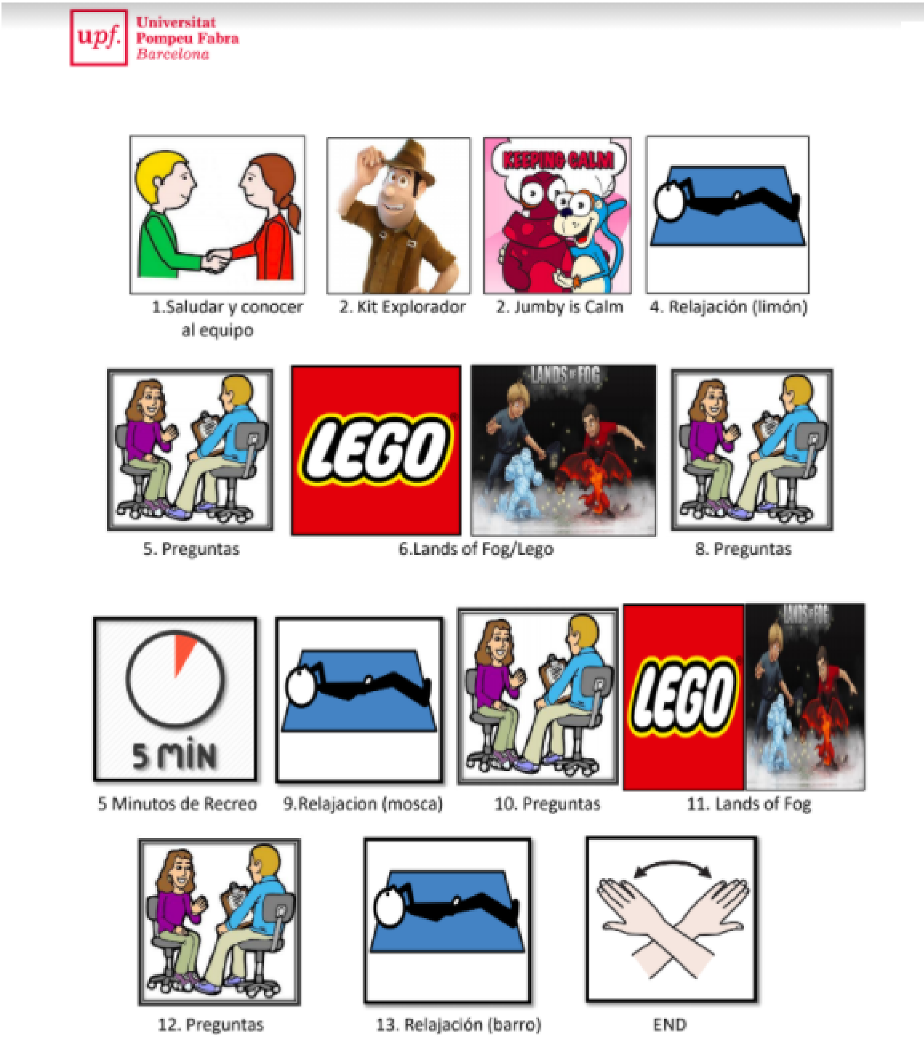


Figure A.2: Visual support schedule of the experimental session

Experiment Date			
Experiment Time			
Starts w LEGO (1) / LOF (2)			
min	sec		
		digital I/O w trigger button	
		digital I/O w trigger button	
		Tablet 1 Sync	
		Tablet 2 Sync	
		relaxation exercise	
		questionnaire_AS (NON-ASC)	
		questionnaire_AS (ASC)	
		NON-ASC baseline (1 min)	
		ASC baseline (1 min)	
		1st Condition (15 min) (LOF or LEGO)	
		questionnaire_AS & questionnaire_STAIC/S & questionnaire_Game (NON-ASC)	
		questionnaire_AS & questionnaire_STAIC/S & questionnaire_Game (ASC)	
		NON-ASC baseline (1 min)	
		ASC baseline (1 min)	
		BREAK (5 min)	
		relaxation exercise	
		questionnaire_AS (NON-ASC)	
		questionnaire_AS (ASC)	
		NON-ASC baseline (1 min)	
		ASC baseline (1 min)	
		2nd Condition (15 min) (LOF or LEGO)	
		questionnaire_AS & questionnaire_STAIC/S & questionnaire_Game (NON-ASC)	
		questionnaire_AS & questionnaire_STAIC/S & questionnaire_Game (ASC)	
		NON-ASC baseline (1 min)	
		ASC baseline (1 min)	
		relaxation exercise	
		FINISH	

Figure A.3: Diary of the experiment

Parc de Salut Barcelona MAR

CONFIRMATION

I hereby certify that “Parc de Salut MAR - Clinical Research Ethics Committee” has reviewed and approved

✓ The project entitled: *“ASCME-OR (Autism Spectrum Condition Multimodal Embodiment Open Repository - A Multimodal On-Line Analysis & Database Tool for Investigating Physiological Effects of Full-body Interaction Virtual Environments for Children with Autism ”*

In which Dr. NARCÍS PARÉS BURGUÈS, is the principal investigator, from the CMTech, Universitat Pompeu Fabra.

✓ And the corresponding participant information sheet and informed consent form.

Under the reference number: 2017/7401/I

At the meeting of: May 23th, 2017

COMITÈ ÈTIC D'INVESTIGACIÓ CLÍNICA
CEIC - PARC DE SALUT MAR

Mª Teresa Navarra Alcrudo
Secretary
Ethics Committee – Parc de Salut MAR

Barcelona, May 24th, 2017

Figure A.4: Ethics committee approval

Tablet - Based Questionnaires: ASCMEOR trials
Pre-condition (questionnaire_AS)
How active do you feel right now?
What is your mood?
How well do you know your partner?
Do you want to know them better?
Post Lands of Fog (questionnaire_AS & questionnaire_STAIC/S & questionnaire_Game)
How active do you feel right now?
What is your mood?
questionnaire_STAIC/S
Do you remember a moment in which you talked or discovered something with your partner?
How active did you feel in this moment?
Did you like it?
What was the most memorable creature of the game?
How active did you feel when this creature appeared?
Did you like it?
What things of the game did you like?
What things did you not like?
In comparison to the schoolyard, would you like to play the videogame with your partner?
How well do you know your partner?
Do you want to know them better?
What do you think this experience was about?
Post LEGO (questionnaire_AS & questionnaire_STAIC/S & questionnaire_Game)
How active do you feel right now?
What is your mood?
questionnaire_STAIC/S
How active did you feel in this moment?
Did you like it?
What things of the game did you like?
What things did you not like?
How well do you know your partner?
Do you want to know them better?
What do you think this experience was about?

Figure A.5: Multitouch tablet questionnaire items