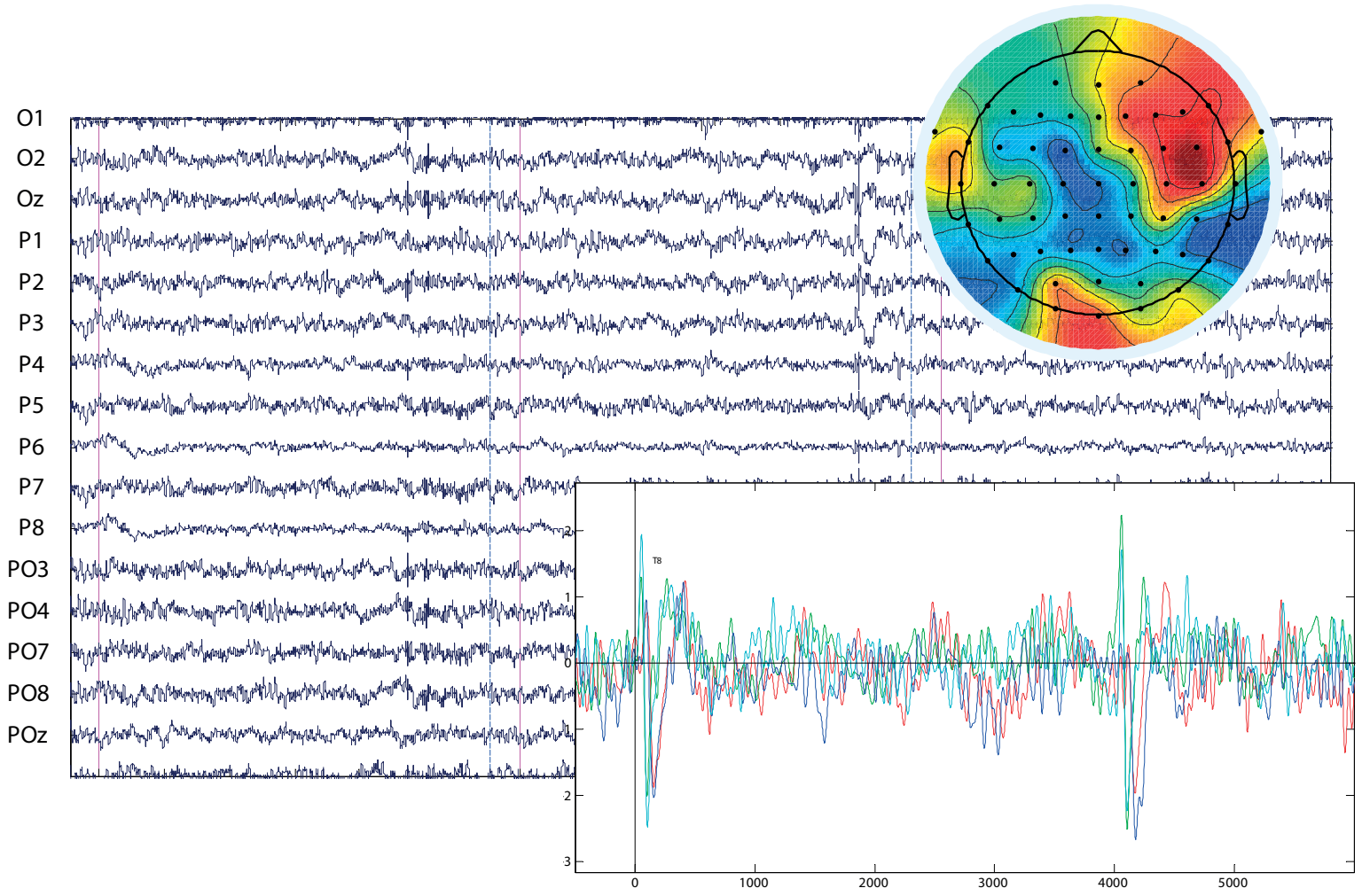

Raquel Vaz Maia Monteiro

BCI APPLICATIONS REGARDING THE PERCEPTION OF EMOTIONS IN HEALTHY INDIVIDUALS AND AUTISM

Dissertation submitted to the University of Coimbra in compliance with the requisites for the degree of Master in Biomedical Engineering

September 2016

· U C ·

Universidade de Coimbra

• U C •

FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Raquel Vaz Maia Monteiro

BCI applications regarding the perception of emotions in healthy individuals and autism

Dissertation submitted to the University of Coimbra in compliance with the requisites for the degree of Master in Biomedical Engineering

Supervisors:

MD PhD Miguel Castelo-Branco (Institute for Biomedical Imaging and Life Sciences, Faculty of Medicine, University of Coimbra)

MSc Marco Simões (Institute for Biomedical Imaging and Life Sciences, Faculty of Medicine, University of Coimbra; Center for Informatics and Systems, University of Coimbra)

MSc João Andrade (Institute for Biomedical Imaging and Life Sciences, Faculty of Medicine, University of Coimbra)

Coimbra, 2016

This work was developed in collaboration with:

Institute for Biomedical Imaging and Life Sciences, Faculty of Medicine, University of Coimbra

Esta cópia da tese é fornecida na condição de que quem a consulta reconhece que os direitos de autor são pertença do autor da tese e que nenhuma citação ou informação obtida a partir dela pode ser publicada sem a referência apropriada.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognize that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledgement.

Aos meus pais,

Acknowledgments

Começo por agradecer, em primeiro lugar, ao Professor Miguel Castelo-Branco, pela oportunidade de trabalhar na “casa” que é o IBILI, por toda a disponibilidade, apoio e supervisão. Agradeço igualmente aos meus dois orientadores, que trabalharam diretamente comigo, Marco Simões e João Andrade. Agradeço-lhes por todo o acompanhamento, apoio e orientação durante todo trabalho, e também pela amizade que fica.

Agradeço igualmente a todos os outros elementos do grupo do Professor Miguel Castelo-Branco, no IBILI, pela forma como me receberam e integraram. Um agradecimento especial aos elementos da “Sala de Cálculo”, por toda amizade. Agradeço por todas as horas de almoço e lanche inesquecíveis, jogos, brincadeiras e debates, por todos os momentos de escape que permitiram aliviar o stress e limpar a alma nos dias mais difíceis.

Como a família é sem dúvida o nosso maior suporte, devo um agradecimento muito especial à minha mãe e ao meu pai por todo o esforço, dedicação, compreensão, afeto e incentivo ao longo destes cinco anos. Agradeço igualmente aos meus avós, maternos e paternos, por todo o carinho e incentivo. Aos meus familiares mais próximos, por toda a preocupação e carinho. À Maria João, à Andreia, à Maria dos Anjos e à Beta um enorme obrigada, por terem zelado pelo meu bem-estar quando eu mais precisei.

À Dr.ª Ana Melo, um sincero obrigada, pois tudo seria mais difícil sem ela. Obrigada por nunca me ter deixado cair.

Um enorme obrigada aos meus amigos mais especiais Rute Marto, Sara Oliveira, Carolina Travassos, Mafalda Lima, Filipa Rodrigues, Mariana Barros Alexandre Sayal, João Marques, André Pedrosa, André Matos e Fábio Reis, por todos os momentos inesquecíveis ao longo destes cinco anos, por toda a amizade e companheirismo.

Por último, à Universidade de Coimbra, ao Departamento de Física e a todos os seus funcionários um grande obrigado pela forma como me acolheram durante estes cinco anos memoráveis.

Resumo

As relações sociais são essenciais para uma vida saudável, evitando o isolamento social que pode conduzir a condições de ansiedade clinicamente relevante e, até mesmo, de depressão. As expressões faciais funcionam como sinais sociais, dando informação sobre o estado emocional do “outro” e, por isso, constituem a chave para a comunicação social. Desde o primeiro ano de vida que as crianças conseguem reconhecer sinais sociais a partir das faces. A capacidade de interpretar corretamente esses sinais é crítica para se ser bem-sucedido no desenvolvimento de interações diádicas ou multidiádicas no dia-a-dia. Contudo, os défices no domínio social constituem o défice mais importante na Perturbação do Espectro do Autismo (PEA). PEA é uma patologia do neuro-desenvolvimento caracterizada por défices na comunicação e interação social, bem como um padrão de comportamentos repetitivos e interesses restritos. A literatura existente em estudos com eletroencefalografia (EEG) reporta diferenças de grupo entre indivíduos com PEA e indivíduos controlo durante tarefas que envolvem reconhecimento e processamento de expressões faciais.

O principal objetivo do trabalho desenvolvido consistiu em avaliar a viabilidade do desenvolvimento de uma aplicação de *neurofeedback,* com recurso a EEG, para reabilitar indivíduos com PEA no que diz respeito aos seus défices a nível social, através do reconhecimento e imaginação do “outro” a fazer expressões faciais. Para tal, foi adquirido sinal EEG de dezassete adolescentes com PEA e dezassete indivíduos controlo durante uma tarefa com uma primeira parte de estimulação visual com expressões faciais, seguida de uma parte de imaginação do “outro” a realizar a expressão facial da instrução. Usando os dados da imaginação de expressões faciais foram extraídas *features* do domínio temporal e *features* não lineares com o objetivo de classificar entre estados de imaginação e não-imaginação. A classificação foi realizada recorrendo a *Support Vector Machines*. Os resultados de classificação obtidos não foram, porém,

promissores no que diz respeito ao desenvolvimento da aplicação de *neurofeedback* para reabilitação social dos indivíduos com PEA.

Como análise complementar avaliou-se a existência de um possível biomarcador eletrofisiológico que permitisse distinguir os indivíduos dos dois grupos, durante o período de imaginação. Usando as mesmas *features* de domínio temporal e não lineares, através de uma estratégia de classificação *leave-one-out*, cada participante do estudo foi atribuído a um dos dois grupos. Os resultados indicam que foi possível distinguir entre os dois grupos com uma precisão entre 80 % a 90 %. As *features* mais relevantes para este processo de classificação revelaram ser das bandas de frequência teta, beta e gama e, maioritariamente determinadas das regiões temporo-parietais e parieto-occipitais do hemisfério direito. Estas regiões correspondem também à localização das estruturas cerebrais responsáveis pelo processamento da face.

Palavras-chave: eletroencefalografia (EEG), Perturbação do Espectro do Autismo (PEA) processamento de expressões faciais, emoções, imaginação.

Abstract

Social relationships are essential for a healthy life, preventing social isolation, which could lead to anxiety and depression. Faces are the key to social communication, conveying information about the emotional state of a person. Since the first year of life, young children are capable of understanding and processing facial signals, like facial expressions. The ability to interpret these social signs is critical for the development of social relationships and successful social interactions. However, deficits in social domain are considered the hallmark domain of impairment in Autism Spectrum Disorder (ASD). ASD is a neurodevelopmental disorder characterized by difficulties in social communication and interaction, and a pattern of restricted and repetitive behaviors and interests. Previous electroencephalography (EEG) studies reported differences in processing and recognition of emotional facial expressions between ASD and typically developing age-matched individuals.

The primary goal of this work was to evaluate the viability of an EEG-based neurofeedback approach to rehabilitate social-emotional reciprocity impairments in ASD through the recognition of facial expressions and mental imagery of a third person performing it. To achieve this goal, we recruited seventeen male teenagers with ASD and seventeen typically developing male teenagers for this study. We acquired EEG signals during a visual stimulation and a mental imagery task of a third person performing happy and sad facial expressions. From the mental imagery task, we extracted time-domain and non-linear-domain features to compare between two conditions (mental imagery of emotional facial expressions and no-imagery states). We developed a classification platform based on features extracted using Support Vector Machines for predicting, for each subject, between the two conditions. Results obtained could not separate between the two conditions with sufficiently high accuracy, which precludes at this stage the development of the EEG neurofeedback application for the rehabilitation of social cognition impairments in ASD individuals.

As a complementary analysis, we tried to find an electrophysiological biomarker for ASD using the mental imagery period. The same time-domain and non-linear features were extracted for each subject and, using a leave-one-out approach, each subject was associated to one of the groups. We were able to separate the two groups with an accuracy value between 80 % and 90 %. The most relevant features were extracted from theta, high-beta and gamma frequency bands, mostly from temporo-parietal and parieto-occipital areas from the right hemisphere, which corresponds to the localization of the brain structures responsible for face processing.

Keywords: electroencephalography (EEG), Autism Spectrum Disorder (ASD), facial expressions processing, emotions, mental imagery.

List of Acronyms

ADHD	Attention Deficit / Hyperactivity Disorder
ADI-R	Autism Diagnostic Interview-Revised
ADOS	Autism Diagnostic Observation Schedule
ALS	Amyotrophic Lateral Sclerosis
APA	American Psychiatric Association
ApEn	Approximate Entropy
ASD	Autism Spectrum Disorder
BCI	Brain Computer Interface
CNS	Central Nervous System
CorrDim	Correlation Dimension
CSP	Common Spatial Pattern
DMS-5	Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
EEG	Electroencephalography
Env	Hilbert Envelope
EOG	Electrooculography
ERP	Event-Related Potential
FFA	Fusiform Face Area
fMRI	Functional Magnetic Resonance Imaging
FS	Fisher Score
ICA	Independent Component Analysis
IQ	Intelligence Quotient
LPA	Left Pre-Auricular Point
Lyap	Largest Lyapunov Exponent
OFA	Occipital Face Area
PCA	Principal Component Analysis
Pow	Power

PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
pSTS	Posterior Part of the Superior Temporal Sulcus
RPA	Right Pre-Auricular Point
SCP	Slow Cortical Potential
SD	Standard Deviation
SEM	Standard Error of the Mean
SFI	Spatial Filling Index
SpEn	Sample Entropy
STS	Superior Temporal Sulcus
SVM	Support Vector Machines
TD	Typically Developing
Teag	Teager Energy
VR	Virtual Reality
WAIS	Wechsler Adult Intelligence Scale
WISC	Wechsler Intelligence Scale for Children

List of Tables

List of Figures

Contents

Chapter 1

Introduction

Faces are the most important visual stimuli we perceive, conveying information about identity and emotional state of a person (Kanwisher and Yovel, 2006). Since the first year of life young children are capable of understanding and processing facial cues, like facial expressions (Bayless et al., 2011). The ability to interpret these social signs represents an essential skill in child development and, therefore, a basic condition for the development of social relationships and successful social interactions early in life (Bayless et al., 2011; Berggren et al., 2016). Therefore, deficits in the social domain lead to relevant medical conditions and are considered the hallmark of Autism Spectrum Disorder (ASD), particularly facial emotion recognition (Crider and Pillai, 2016).

ASD is a neurodevelopmental disorder characterized by deficits in social communication and interaction, and a pattern of restricted and repetitive interests, behaviors or activities (American Psychiatric Association, 2013b). In recent years the number of individuals with ASD reached 1% of the population (American Psychiatric Association, 2013a; Centers for Disease Control and Prevention, 2016).

Social impairments are frequently debilitating and can diminish quality of life, leading to social isolation and comorbid conditions like anxiety and depression; thus, it is important to understand the nature of that impairment and to develop strategies of intervention that enable social rehabilitation of these individuals. As ASD individuals show a strong motivation towards technologies, the use of innovative approaches to rehabilitate these individuals, like neurofeedback and Brain Computer Interfaces (BCI), can be particularly effective when compared to behavior and pharmacological therapies (Coben, Linden, and Myers, 2010; Wainer and Ingersoll, 2011).

1.1 Problem Definition

The aim of the project is to evaluate the viability of the development of an electroencephalography (EEG)-based BCI system, with neurofeedback, to rehabilitate social impairments in individuals with ASD. A BCI system is a direct communication pathway between the brain and an external device, and neurofeedback is the concept of providing real-time information to an individual based on their own brain activity, enabling him to learn how to self-regulate it. A deeper explanation of these methods can be found on Chapter 2. The rehabilitation process will focus on the facial emotional expression recognition impairments in these individuals.

The IBILI research group has recently developed a similar neurofeedback approach using functional Magnetic Resonance Imaging (fMRI). This tool was based on activation of the posterior part of the Superior Temporal Sulcus (pSTS) during a task involving the processing and recognition of facial emotional expressions, and imagery of a third person performing facial expressions. As EEG has better temporal resolution than hemodynamic measures (fMRI) and it is less expensive, there is high interest in transferring the knowledge obtained in fMRI to EEG.

To achieve this transfer to the EEG technique, we investigated brain responses to dynamic facial expression stimuli, for both ASD and healthy individuals. Additionally, we searched for neural correlates of imagery of a third person performing facial expressions, and assessed the viability of an EEG-based BCI with neurofeedback, based on the correlates found.

This EEG-based tool aims to train the user in two specific tasks of social cognition:

- recognize emotional facial expressions performed by a third person,
- understand the other's emotional state based on the facial expression performed, enabling these individuals to imagine the third person performing facial expressions.

1.2 Objectives

We aimed to:

- perform a literature systematic review on facial expression processing in ASD individuals using EEG technology,
- study the neural correlates of dynamic response to facial expression stimuli and mental imagery of a third person performing facial expressions (involving data acquisition with ASD teenagers and age- and Intelligence Quotient (IQ)-matched healthy controls, and data analysis with existing methods and software tools),
- extraction of EEG signal features associated to mental imagery of a third person performing facial expressions,
- evaluate the viability of an EEG-based BCI with neurofeedback approach based on the identified features.

This project is part of the work performed within the scope of a doctoral thesis which aims to evaluate the methods to do the transfer between fMRI and EEG. Thus, the data acquisition and analysis were done in collaboration with Marco Simões and João Andrade, while literature review, feature extraction and evaluation of viability of the EEG-based BCI/neurofeedback approach were done particularly by me with their supervision; thus, this dissertation is primarily focused on that part. This work was developed under the guidance and supervision of Professor Miguel Castelo-Branco.

1.3 Structure of the Dissertation

Six chapters constitute this dissertation. The first chapter is the introduction, where the problem was defined with its motivation, and the objectives and main contributions of this work were stated. The second chapter contains a theoretical background where it was explored the clinical context of the Autism Spectrum Disorder, with special focus to social-emotional reciprocity impairments, and

theoretical concepts important for the comprehension of the following contents. The third chapter presents the state of the art of emotional facial expression processing in ASD using electrophysiological measures, and EEG-based BCI and neurofeedback approaches for individuals with ASD. In the fourth chapter, it is described the experimental task used to acquire EEG data on ASD and typically developing individuals, the methods used to preprocess and analyze the data, feature extraction and evaluation of the viability of a EEG-based BCI with neurofeedback approach to rehabilitate social cognition in these individuals. The fifth chapter presents the results obtained, and the sixth chapter presents the discussion and conclusions of the work.

Chapter 2

Theoretical Background

2.1 Autism Spectrum Disorder

Autism Spectrum Disorder (ASD) was the name adopted in 2013 by American Psychiatric Association (APA), in the 5th edition of Diagnostic and Statistical Manual of Mental Disorders (DSM-5), for a group of four previously separate disorders:

- Autism,
- Asperger's Syndrome (a milder form of autism),
- Childhood Disintegrative Disorder,
- Pervasive Developmental Disorder Not Otherwise Specified (or atypical autism).

As the word "spectrum" suggests, ASD is a heterogeneous, behaviorally defined neurodevelopmental disorder, with multiple causes and courses, and a great range of severity and symptoms (Amaral et al., 2008; Frith et al., 2005). Autism is the major disorder of the spectrum and, therefore, the other disorders of the spectrum were not detailed since they are not fundamental for the comprehension of the project and its purpose.

2.1.1 Clinical Characterization

Neurodevelopmental disorders are a group of conditions characterized by deficits, with onset in the developmental period, that produce impairments in personal, social and academic functioning (American Psychiatric Association, 2013a).

“Autistic Disturbances of Affective Contact” (Kanner, 1943), published in 1943 by Dr. Leo Kanner, an Austrian psychiatrist, was the first clinical account of the disorder. In his study, based on direct observation of 11 children, he described essential features of autism, all of which are still valid in diagnostic manuals nowadays. Dr. Kanner (Kanner, 1943) described a group of highly intelligent children who displayed a strong desire for aloneness and an inability to relate with others, from the beginning of life. Marked limitation on the variety of spontaneous activities, repetitive behavior and obsessive desire for the maintenance of sameness were also typical behaviors. Currently, according to DSM-5, ASD is described as a neurodevelopmental disorder characterized by difficulties in social communication and social interaction, and a pattern of restricted and repetitive interests, behaviors, or activities (American Psychiatric Association, 2013b). Table 2.1 describes the main impairments of ASD according to DSM-5. In adults with autism which do not present intellectual impairments or language delays, deficits in social-emotional reciprocity may be most apparent impairments in processing and responding to complex social cues.

Studies show that the majority of adults with autism have poor outcomes in terms of independent living, employment and peer relationships. Therefore, early behavioral interventions could help to improve social functioning and reduce anxiety and aggressiveness in adulthood (Lai et al., 2014).

Table 2.1 – Main impairments of Autism Spectrum Disorder according to DSM-5 (American Psychiatric Association, 2013a).

Persistent deficits in social communication and social interaction	Deficits in social-emotional reciprocity	Abnormal social approach
		Reduced sharing of interests, emotions or affects
		Failure to initiate or respond to social interactions
	Deficits in non-verbal communicative behaviors	Abnormalities in eye contact and body language
		Lack of facial expressions and nonverbal communication
	Deficits in developing, maintaining and understanding relationships	Difficulties in adjusting behavior to suit various social contexts
		Difficulties in sharing imaginative play or making friends
		Absence of interest in peers
Restricted and repetitive patterns of behavior and interests	Stereotyped or repetitive motor movements, repetitive use of objects or repetitive speech	
	Insistence on sameness, inflexible adherence to routines, or ritualizes patterns of verbal and nonverbal behavior	
	Highly restricted interests, that are abnormal in intensity or focus	
	Hyper- or hypo-reactivity to sensory input or unusual interests in sensory aspects of the environment	

2.1.2 Prevalence and Diagnostic

In recent years, due to monitoring and diagnosis, the number of individuals with ASD reached 1% of the population. Statistics show that the overall prevalence of ASD is 4 to 5 times higher in males compared to females, however empirical data suggests females are diagnosed later than males, indicating a diagnose bias

towards males (American Psychiatric Association, 2013a; Centers for Disease Control and Prevention, 2016; Lai et al., 2014).

Symptoms are typically recognized during the second year of life, although the stage at which functional impairments become obvious varies according to the characteristics of the individual and its environment, and also depends on the severity of the autistic condition, developmental level and chronological age (American Psychiatric Association, 2013a). Primary symptoms of ASD include deficits or delays in joint attention and pretend play, deficits in reciprocal affective behavior (unusual social interaction and lack of social interest), delayed verbal and non-verbal communication, unusual repetitive behaviors and extreme variation in temperament (American Psychiatric Association, 2013a; Lai et al., 2014).

ASD may have several comorbid disorders associated. More than 70% of the cases have concurrent conditions, like intellectual and/or language impairment (American Psychiatric Association, 2013a; Lai et al., 2014). Emotional problems, anxiety and depression, seizure disorders like epilepsy, and Attention Deficit/Hyperactivity Disorder (ADHD) can also be associated with ASD (Amaral et al., 2008; Marzbani et al., 2016). Anxiety and depression can also contribute to social isolation; pharmacological therapies can reduce these comorbid symptoms, but do not directly improve social-communicative deficits (Crider and Pillai, 2016; Lai et al., 2014).

2.1.3 Social-Emotional Reciprocity Impairments

Deficits in social domain are considered the hallmark domain of impairment in ASD, as many of its symptoms contribute to social isolation (Crider and Pillai, 2016). These deficits may not be immediately detected in early years. However, they become gradually more evident as they prevent children to develop normal interpersonal relationships with their family and other children.

Individuals with autism have specific impairments in the processing of social and emotional information, and these deficits can reliably be identified regardless of Intelligence Quotient (IQ) (Foss-Feig et al., 2016; Korkmaz, 2011). Social interaction (joint attention and pretend play), social perception (eye contact and

emotion perception) and social cognitive abilities, such as imitation, empathy and theory of mind[1], are social aspects in which individuals with autism are severely impaired (Bernier et al., 2013; Crider and Pillai, 2016; Foss-Feig et al., 2016; Lai et al., 2014).

2.2 Functional Electroencephalography

The human brain is divided into three main parts: cerebrum, brainstem and cerebellum (Figure 2.1). The cerebrum has a special dominant position as it operates in conscious functions of the nervous system. The functional part of the cerebrum is the cerebral cortex, a relatively thin layer of gray matter covering the outer surface of the cerebrum; it has a dense collection of nerve cells which send and receive sensory information to and from the brain and nervous system. Gray matter contains a predominance of cell bodies and the deeper layers, beneath the cortex, consist of myelinated axons and are called white matter. Cerebral cortex has a highly convoluted surface consisting of gyri (ridges) and sulci (valleys) (Clark, 2009; Nunez and Srinivasan, 2006).

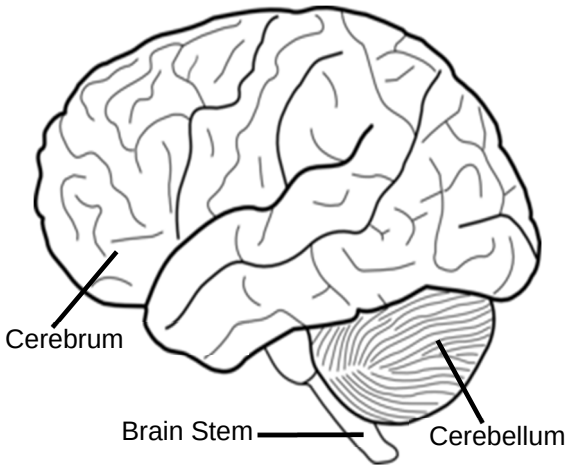

Figure 2.1 – Illustration of the human brain and their main parts.

[1] The ability to conceive mental states of others – understand and infer facial expressions, body language, figurative speech and other social cues that convey emotional information – and understand that these may differ from our own (Dichter, 2012; Gross et al., 2012; Korkmaz, 2011).

Information flow in the brain is due to action potentials, which are small electrical currents that propagate across neurons' membranes (Figure 2.2). An action potential is the result of an increase in the potential inside the axons due to the inflow of positive charges into the cell membrane, which generates electrical impulses that propagate across the axons. When an action potential reaches the end of an axon a set of chemical neurotransmitters is released from the neuron's presynaptic terminal to the synaptic cleft; this process is called synapse. When neurotransmitters bind to postsynaptic receptors (dendrites) a voltage difference arises, called postsynaptic potential, causing the opening or closing of the ion channels in the membrane of the postsynaptic neuron and, thus, leading to an action potential (Luck, 2005).

The synchronized/desynchronized occurrence of multiple potentials in the various regions of the brain can produce a considerable electrical activity which can be measured at cortical surface using electrophysiological measures (Clark, 2009; Dawson et al., 2005).

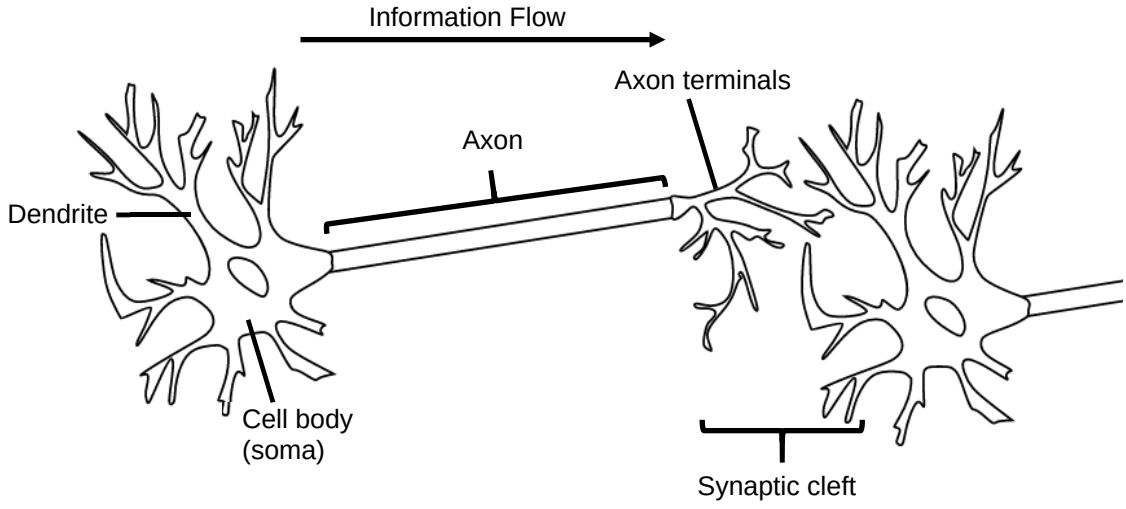

Figure 2.2 – Illustration of the process of flow of information in the brain.

2.2.1 Electroencephalography

An example of a functional electrophysiological technique capable of recording electrical brain activity is the electroencephalography (EEG) which uses scalp electrodes associated with an appropriate recording system (illustrated in Figure 2.3). EEG technique is noninvasive and can be used with participants who have limited cognitive or communicative abilities (Dawson et al., 2005). This technique is widely used in research, and in clinical in the diagnosis of epilepsy, brain death, coma and severe head injury, and monitoring of depth of anesthesia (Dawson et al., 2005; Nazari, 2012; Nunez and Srinivasan, 2006).

An EEG signal is characterized by two major parameters derived from its Fourier components: frequency and amplitude. Frequency indicates how fast the brainwaves oscillate, measured in Hertz, and amplitude represents the power of these waves, measured in microvolts (Marzbani et al., 2016; Nazari, 2012).

To obtain EEG data, one or more electrodes are placed over the scalp with a conductive gel to improve contact of electrode and skin, which diminishes the loss of electrical power of the signal (Davies and Gavin, 2007). The recording locations are most often chosen according to the International 10-20 system (see section 2.2.2), or to its expanded versions, in order to ensure reproducibility (M. Fabiani et al., 2007). Typically, electrodes are embedded in caps or nets in order to facilitate their placement. The electrodes are then connected to an amplifier sensitive to potential differences. Electrical signal is then filtered with anti-aliasing filters and digitalized using an analog-to-digital converter (Davies and Gavin, 2007; M. Fabiani et al., 2007; Guger et al., 2011).

Electrical signals that arise in the EEG signal but are not originated from brain activity are called artifacts, and have to be removed from the signal recorded. These artifacts can be of biological origin, like eye blinks or eye movements, tension of the muscles of the head and neck and electrical activity generated by the heart; or from environmental and experimental origin, like bad grounding (M. Fabiani et al., 2007).

Figure 2.3 – Illustration of an EEG medical setup (Brain Products GmbH, 2016).

An EEG signal is extremely sensitive to real time neural processes, providing detailed temporal resolution on the scale of milliseconds, whereas hemodynamic measures taken from fMRI are limited to a resolution of few seconds. However, these hemodynamic measures have a spatial resolution within the millimeter range, which electromagnetic measures cannot match.

2.2.1.1 The volume conduction problem

Every neuron receiving synaptic inputs can be thought of as a dipole, with specific orientation and polarity. During an action potential, the inflow of positive ions, into the cell membrane, at a given position in the dendrite of a neuron is instantaneously balanced by an outflow of positive charges at another region of the neuron, resulting in a dipole. The instantaneous dipole that appears in each neuron is responsible for the production of a dipole field outside the neuron (Destexhe and Bedard, 2012). When a dipole is present in a conductive medium such as the brain, current is conducted throughout that medium until it reaches the surface (Luck, 2005). This process is called volume conduction. However, in practice, neurological sources do not correspond precisely to one-dimensional dipoles. Any activity detected in the EEG is a result of the summation of many individual dipoles.

The voltage recorded at the scalp surface will depend on the position and orientation of the dipole resultant of the average of the individual dipoles and, also, on the resistance and shape of the various components of the medium which electrical activity has to cross (Clark, 2009; Luck, 2005). For instance, when the signal reaches the skull it tends to spread laterally as it finds a very low conductance medium which smears electric fields, acting as a low-pass spatial filter (M. Fabiani et al., 2007). These factors contribute to the blur of the surface distribution of the voltage, and to hamper the intention to relate the potentials measured on the scalp with their brain current sources and underlying cognitive processes (Luck, 2005). Thus, the reduced spatial resolution is the biggest disadvantage of EEG technique.

2.2.2 International 10-20 System

International 10-20 system is an internationally recognized method that describes the location of EEG scalp electrodes using anatomical landmarks of the skull (Jurcak et al., 2007; Tatum, 2014). This method has been gaining importance as it ensures the standardized reproducibility of EEG experiments. According to this electrode system, the position of the electrodes on the scalp is described via relative distances between cranial landmarks over the head surface. The anatomical landmarks that this system uses are the nasion, point between forehead and the nose; the inion, the external occipital protuberance; the left pre-auricular point (LPA) and the right pre-auricular point (RPA), which are felt as depressions anterior to the ear (Jurcak et al., 2007; Marzbani et al., 2016). The skull perimeters are measured from these four landmarks, in transverse and medial planes. These perimeters are then subdivided by intervals of 10% and 20% where electrodes are placed (Jurcak et al., 2007).

In this system, the skull regions are named using letters, representing the proximity of the electrode with a specific brain region, and numbers, which represent the location in either right or left hemisphere. The designations Fp, F, T, O, C and P represent, respectively, frontopolar, frontal, temporal, occipital, central and parietal brain regions. Subsequently numbers combined following the letters

represent the left hemisphere (odd numbers) and the right hemisphere (even numbers). The letter “z” is used to designate scalp locations that fall along the central midline, running between the nasion and the inion. The numbers reflect the location proximity with the midline (Coles and Rugg, 1996; Marzbani et al., 2016; Tatum, 2014).

The International 10-20 system only allows a small number of EEG electrodes, typically 21. With the development of EEG technique, signal source localization methods become necessary and, with them, a system with higher density of electrodes. Therefore, it was proposed a modified combinatorial electrode system in which electrode placement is more closely spaced providing a higher density of scalp electrode (typically, 81), named International 10-10 system (Jurcak et al., 2007; Tatum, 2014). The International 10-20 system does not name the intermediate electrode positions created in the International 10-10 system so, this combinatorial system designates them by combining the letters for the two standard electrode positions that surround them (American Clinical Neurophysiology Society, 2006). Thus, frontotemporal (FT) and frontocentral (FC) are the designations of electrode positions in the second intermediate coronal line; temporo-parietal (TP) and centro-parietal (CP) in the third intermediate coronal line; and, anterior-frontal (AF) and parieto-occipital (PO) are the designations of the positions in the first and last intermediate coronal line, respectively (American Clinical Neurophysiology Society, 2006). In the 10-10 system some positions change its designation, like T3/T4 which become T7/T8 and T5/T6 which become P7/P8 (Tatum, 2014). Figure 2.4 illustrates the electrode positions according to the International 10-20 system (left) and to the International 10-10 system (right).

However, for studies on topography and source analysis of EEG activity, the scientists sought for even higher resolution systems. Thus, Oostenveld and Praamstra (2001) logically extended the International 10-10 system to the International 10-5 system, an even higher density electrode setting. This system enables the use of more than 300 electrode locations (Jurcak et al., 2007).

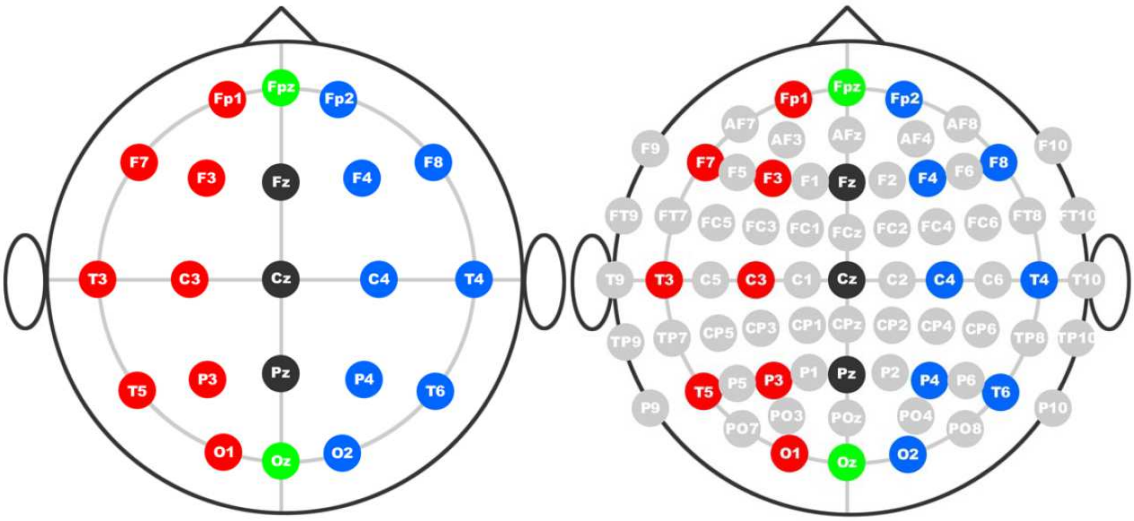

Figure 2.4 – Illustration of electrode positions in the International 10-20 System (left) and International 10-10 System, or Extended 10-20 System (right). Adapted from (Trans Cranial Technologies ldt., 2012).

2.2.3 EEG Rhythms

An EEG signal recorded between electrodes placed on the scalp consists of brain oscillations with different characteristics. These different patterns of electrical activity consist on a brainwave that varies in time, as such they can be recognized by their amplitudes and frequencies.

Rhythmic activity within a specific frequency range was noted to have a certain distribution over the scalp or was associated with certain states of brain function. Therefore, waveforms were subdivided into bandwidths known as delta, theta, alpha, beta and gamma. Frequency bands are usually extracted using spectral methods, like Fourier Transform. Although there is no agreement on what the values of the frequency ranges should be (Marzbani et al., 2016), they are more or less standardized.

In general, the frequency and amplitude of brain oscillations are negatively correlated. As amplitude of the oscillations is proportional to the number of synchronously active neurons, this means that slow brain waves comprise more neurons than fast oscillating brain waves (Pfurtscheller and Lopes, 1999).

2.2.3.1 Delta Rhythm

Delta waves include all the waves in the EEG signal with frequencies below 4 Hz. These rhythms occur in deep sleep. In waking states, delta rhythms are generally abnormal and they can be focal (subcortical lesions) or diffuse (generalized dysfunction). It is predominantly frontal in adults and posterior in children (Clark, 2009; Medscape, 2014; Tatum, 2014).

2.2.3.2 Theta Rhythm

Theta rhythms are composed of signals with frequencies between 4 and 8 Hz, and varying amplitude and morphologies. These waves are normally seen in young children, or during sleep. The appearance of frontal theta can be facilitated by emotions, focused concentration and during mental tasks (Tatum, 2014). Usually associated with creative states, fantasy, imagery, dreamlike and drowsiness (Collura, 1997).

2.2.3.3 Alpha Rhythm

Alpha waves occur at a frequency of 8-15 Hz. They are observed when a person is relaxed and awake, but with eyes closed. This rhythm disappears when subjects are asleep and when subjects' attention is directed to some specific type of mental activity. In normal EEG, the largest contributions to the alpha rhythm come from occipital and parietal regions with somewhat lower contributions from the frontal regions (Clark, 2009; Nunez and Srinivasan, 2006; Tatum, 2014).

2.2.3.4 Beta Rhythm

Beta waves normally occur in the frequency range of 15 to 30 Hz. Beta rhythms are observed when a person is active and alert, with anxious thinking and active concentration. It is the dominant rhythm in patients who are alert or have their eyes open. These waves are most frequently recorded from the parietal and

frontal regions of the scalp, and its distribution on both hemispheres is quite symmetrical. Many authors divide this very large frequency band in sub bands, named low and high beta, since they found that these different sub bands of the same rhythm exhibit different behaviors and performances. Frequency ranges of these sub bands are not well established as they vary between studies (Clark, 2009; Tatum, 2014).

2.2.3.5 Gamma Rhythm

Gamma rhythm is composed of frequencies higher than 30 Hz. These waves are very localized and are observed when a person is processing high-level information, for example when a person is trying to solve a problem and learning.

2.2.4 Event-Related Potentials

Electrical brain activity can be divided into two main categories: spontaneous potentials, like sleep rhythms, and evoked or event-related potentials (ERPs). Evoked potentials are the direct response to an external stimulus, like a flashlight. ERPs are small changes in electrical brain activity that result from synchronized activation of population of neurons in response to some internal or external event. The ERP signal is caused by high cognitive processes that might involve memory updating, attention, expectation, semantic comprehension or changes in mental states, activated in response to a specific stimulus (M. Fabiani et al., 2007; Luckhardt et al., 2014; Otten and Rugg, 2005).

The ERP signal is small in comparison to the EEG signal, because EEG signal reflects thousands of simultaneously ongoing brain processes and the ERP reflects the brain response to a single stimulus of interest. For that reason, ERP signal is not usually visible in a single sample of the EEG recording. The most common way to overcome this and obtain the brain response to the stimulus is to average a large number of samples of the EEG that are time-locked to the repeated occurrences of the particular stimulus, named trials. The averaging procedure acts

as a low-pass filter and enhances signal-to-noise ratio, causing spontaneous potentials of the EEG signal to be averaged out and the relevant ERP signal to remain (M. Fabiani et al., 2007). However, this simple and widely used procedure is based on the assumption that the ERP signal is invariant across trials and, thus it reveals a brain signal that is just an approximation to the reality of the brain response to the stimulus. (Otten and Rugg, 2005; Pfurtscheller and Lopes, 1999).

Event-related signal consists of a waveform containing a series of characteristic peaks, named components, and may be related with specific cognitive processes (ERP) or with no task (evoked potentials) (M. Fabiani et al., 2007; Nunez and Srinivasan, 2006). ERP components are commonly described by its amplitude, polarity, latency, scalp distribution and its relation to experimental variables. The amplitude of the component indicates the amount of allocation of neural resources to specific cognitive processes, and the latency informs about the time course of the processing activity in milliseconds (Duncan et al., 2009). ERP components are commonly referred by a letter indicating the polarity of the peak amplitude (P – positive, N – negative), followed by a number indicating the latency in milliseconds or the number of the ordinal position of the component in the waveform.

Initially, researchers focused their interest on spontaneous rhythmic activity. However, in recent years, researchers have been interested in understanding how the brain responds to specific cognitive functions that are activated in response to a specific time-locked event/stimulus in an epoch of the EEG signal, i.e., ERPs. An advantage of this technique is its high temporal resolution which permits making inferences about the timing of cognitive processes (M. Fabiani et al., 2007; Otten and Rugg, 2005). ERPs are specific to spatial regions of the brain; therefore, ERP signal can map the processing of a visual stimulus from the first processing stages in the primary visual cortex to later cognitive stages such as face identification. The ERP signals resultant of visual stimuli can be grouped in a subset named Visual Evoked Potential (VEP). P100 is one of the earliest visual components, it represents the perception and other early stages in visual processing; N170 is an ERP component that is related with face processing, and it is considered face sensitive and responds to the structural processing of faces (Luckhardt et al., 2014). Figure 2.5 illustrates these two ERP components.

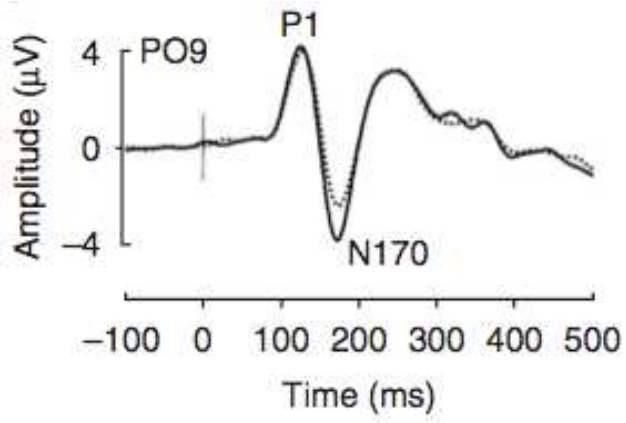

Figure 2.5 – Illustration of P1 and N170 ERP components of an EEG signal. Adapted from (Johan, 2007).

2.3 Neurofeedback and Brain Computer Interfaces

2.3.1 Neurofeedback

Biofeedback is a process that enables an individual to learn how to change his physiological activity for the purposes of improving health and performance (Association for Applied Psychophysiology and Biofeedback, 2011). Neurofeedback is a type of biofeedback that uses different techniques, like EEG, to provide real-time information of brain activity to an individual, enabling him to learn how to self-regulate his brain oscillations *in vivo* and in near real-time (Friedrich et al., 2015; Nazari, 2012). The possibility of volitional control of brain electrical activity suggests that it is related to cognitive functions, and its modulation can have a functional impact (Friedrich et al., 2015). During neurofeedback training, brain activity is recorded with EEG, for example. Then, the electrical signal recorded is amplified and preprocessed using specific software. Different features of the signal are extracted and fed back to the subject thus creating the online feedback loop. During this procedure the individual becomes aware of the changes in his brain activity and tries to improve his brain patterns in order to achieve

optimum performance (Marzbani et al., 2016; Nazari, 2012). Training systems usually provide visual feedback, auditory feedback or verbal feedback given by the trainer (Collura, 1997).

Neurofeedback is a safe and non-invasive technique that is used in basic and applied neuroscience as well as in clinical practice as an alternative choice of treatment (Q. Wang et al., 2010). Neurofeedback techniques can be used to train EEG rhythms and it is claimed that self-regulation of brain electrical activity result in a therapeutic benefit (Marzbani et al., 2016; Nazari, 2012). The application of this technique in clinical practice showed improvement in many disorders like ADHD (Bluschke et al., 2016; Deilami et al., 2016), anxiety and depression (Cheon, Koo, and Choi, 2016), autism (Coben et al., 2010; Friedrich et al., 2015; Kouijzer et al., 2013), epilepsy (Strehl et al., 2014), insomnia (Schabus et al., 2014), learning disabilities (Fernández et al., 2016), dyslexia (Nazari et al., 2012) and schizophrenia (Surmeli et al., 2012). However, its validity in conclusive scientific evidence and its effectiveness had been questioned; it is also time-consuming and its benefits are not long-lasting (Marzbani et al., 2016).

2.3.2 Brain Computer Interfaces

An EEG-based Brain Computer Interface (BCI) is a communication system between an external controlling device and the human brain. These systems are widely used in clinical and research applications, and allow people to communicate through direct measures of the brain activity, without requiring any movement (Cecotti and Gräser, 2011). Nowadays, many universities and laboratories are trying to provide more interactions with virtual reality through BCI (Marzbani et al., 2016; Ploog et al., 2013; Wainer and Ingersoll, 2011).

BCI systems can be invasive or non-invasive. Non-invasive BCI systems use, in most of the cases, EEG. Invasive BCI systems, in which electrodes are placed directly on the cortex, do not have the same disadvantages as EEG-based systems do, like signal blurring by the skull. However, since the electrode placement requires surgery, the range of applications is considerably narrower.

BCI has applications in rehabilitation, neuroscience and cognitive psychology (Jue Wang et al., 2007). Existing research in applications of BCI is composed of

two main areas: for assistive technology, helping disabled people to communicate through and with machines, and as a therapy tool. The primary use of this technology is to benefit people with blocking diseases, such as Amyotrophic Lateral Sclerosis (ALS), brainstem stroke, cerebral palsy, or people with motor disabilities whom have suffered some kind of traumatic accident. This technology makes it possible for these people to regain interaction with external environment in order to improve their quality of life by controlling, with their brain signals, computers, wheelchairs, prostheses, robotic systems and other devices (Figueiredo et al., 2011; Martínez and Barrientos, 2011; Jue Wang et al., 2007). As a therapy tool, BCI systems helps subjects to recover their cognitive function by consciously alter some features of their brain signals. There also exist BCI applications for video game controllers and neurofeedback games, presenting real-time feedback to the user based on EEG signals in the form of video and sound display. These “serious” games usually have educational or health-related aims beside entertainment (Q. Wang et al., 2010).

In BCI systems, electrical signals are recorded just like in an EEG. Then, the features extracted are processed in order to detect a specific event, or to identify and recognize cerebral patterns that are going to be used as inputs. Inputs are translated to control commands of the external device (Martínez and Barrientos, 2011). The real-time system has to work fast enough to present feedback to the subject via the stimulation unit (Guger et al., 2011).

EEG classification strategy depends on the stimulus and, thereby, the response to detect. Current BCI systems are mainly based on four different neuro-mechanisms: slow cortical potentials (SCPs), ERPs (like P300 response), evoked potentials or spontaneous EEG related to motor imagery tasks (Cecotti and Gräser, 2011; Guger et al., 2011; Jue Wang et al., 2007). Devices based on evoked potentials or P300 response extract the intention of the users by detecting which target users are gazing at (Yoshimura and Itakura, 2011). A largely studied paradigm based on the P300 response is the BCI speller. Every time the desired character flashes a P300 wave occurs and the system identifies the letter the user wants to choose. In SCPs and motor imagery BCI systems the user has to acquire control of his/her brain rhythms (Guger et al., 2011). Frequency training is the most prevalent method in clinical applications; this method uses spectral analysis and focus on specific frequencies at specific scalp locations (G. E. Fabiani et al., 2004).

Chapter 3

State of the Art

3.1 Facial Expression Processing in Autism Spectrum Disorder

Since the first year of life, young children are capable of understand and process facial signals like direction of gaze and facial expressions. The ability to interpret these social signs is critical for the development of social relationships and successful social interaction (Bayless et al., 2011; Berggren et al., 2016; Dawson et al., 2005). However, many of the early social cognitive impairments in ASD, such as eye contact, joint attention, and face and emotion recognition, involve the ability to process information from faces (Crider and Pillai, 2016; Dichter, 2012; Eack et al., 2013; Harms et al., 2010). Therefore, it is important to understand the nature of those impairments and to develop strategies of intervention that could enable social rehabilitation of these individuals.

Several behavioral studies, but not all, have described impairments in face processing abilities in individuals with autism; in particular, in tasks that involve the emotion processing (Feuerriegel et al., 2015; J. A. Walsh et al., 2016). Eack et al. (2013) examined the error patterns in response to an emotion recognition task in a group of adults with ASD and age- and gender-matched controls. The authors reported significant impairments in the accuracy and speed with which ASD individuals were able to identify emotional and neutral expressions correctly, compared to typical individuals. ASD individuals more frequently misinterpreted happy faces as neutral faces, and confused neutral faces with negative facial expressions (sad and angry). Walsh et al. (2016) compared the performance of a group of adults with ASD and matched typical individuals on four face perception

tasks, that involved identity, basic and complex emotional expression and trustworthiness perception. The ASD group showed poorer performance in the recognition of both simple and complex emotional facial expression recognition tasks; suggesting that face processing impairments in these individuals result from deficits related to processing emotional information rather than a global face processing deficit. Berggren et al. (2016) examined facial emotion recognition with face and eyes stimuli in well-matched samples of ASD, ADHD, and typically developing (TD) individuals. Results are in agreement with the literature, suggesting facial emotion recognition deficits in ASD concerning the response time and the correct identification of the emotion, when compared to TD individuals. However, no differences were found between ADHD and TD individuals.

3.1.1 Neural Systems for Facial Expression Processing

Neuroimaging studies have revealed multiple cortical regions along occipito-temporal cortex that form a neural network specialized in the analysis of faces (Bernstein and Yovel, 2015). In healthy individuals, the core brain regions responsible for face processing include the inferior occipital gyri (occipital face area – OFA), the lateral part of the fusiform gyrus (fusiform face area – FFA) and posterior part of the superior temporal sulcus (pSTS) (Bernstein and Yovel, 2015; Harms et al., 2010). In 2000, Haxby et al. (2000) proposed a model to describe the functional role of this face-selective network. Figure 3.1 illustrates the location of the core brain regions that compose this face-selective network, according to the Haxby model. According to this model, OFA is responsible for early stages of face processing and sends its outputs both to FFA and pSTS, forming two different neural pathways. Invariant aspects of the face, such as identity, involve greater activation of the FFA in the ventral stream. The variable aspects of the faces, such as eye gaze and facial expressions, recruit the pSTS, forming the dorsal stream (Bernstein and Yovel, 2015; Harms et al., 2010; Haxby et al., 2000). The processing of facial emotions also elicits activity in the limbic regions, such as the

amygdala, involved in recognizing emotions from facial expressions, and the insula (Harms et al., 2010; Kashihara, 2014). Several studies have identified these brain structures to be hypoactive in ASD across tasks using social perception and cognition (Coben et al., 2010; Dichter, 2012; Foss-Feig et al., 2016; Lai et al., 2014).

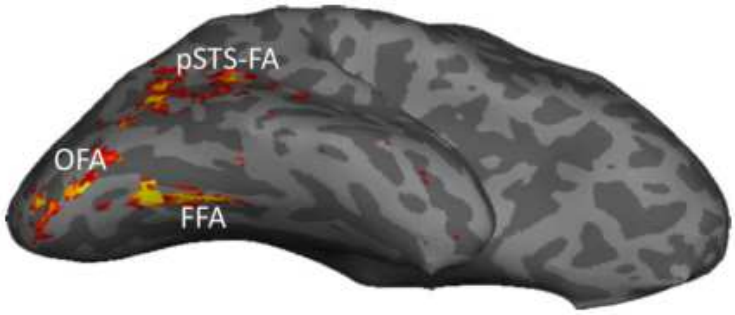

Figure 3.1 – Representation of the core brain regions responsible for face processing, according to the Haxby model. Adapted from (Bernstein and Yovel, 2015).

3.1.2 EEG Studies of Facial Expression Processing in Autism Spectrum Disorder

A systematic review of the existing literature was written in order to evaluate evidence for differences in emotional facial expression processing between ASD (and Asperger Syndrome, to include pre-DSM-5 nomenclature) and TD individuals. A total of 15 articles met the inclusion criteria and were selected for reviewing. These articles investigate impairments in emotion perception in ASD using EEG recorded during the presentation of emotional face stimuli of an unfamiliar person. The most frequent measured ERP components were N170 and P1 recorded over posterior temporal and occipital scalp areas, respectively. These potentials are associated with structural encoding of faces and early low-level stage of holistic face perception, respectively.

Statistically significant differences between groups were found in P1 and N170 amplitudes and latencies. Individuals with ASD showed higher latencies, as well as lower amplitudes, to emotional stimuli (Apicella et al., 2013; Batty et al., 2011; O'Connor et al., 2005; Tye et al., 2014), suggesting some disruption in the processes involved in emotional face perception in ASD (Batty et al., 2011). A main effect of age was verified in two studies (Batty et al., 2011; O'Connor et al., 2005), indicating that P1 and N170 amplitude and latencies both decreased with age. O'Connor and colleagues (2005) found that, in contrast to adults, N170 amplitude and latency differences were not observed between ASD and TD children, which raises the point of incomplete development of the N170 component in children. The authors suggested that N170 component, although present in ASD and TD children, continues to mature into adulthood, since group differences were reported only in adults.

Statistically significant effects were found for emotion and emotion by group interaction. A main effect of emotion, for both groups, was found for the happy facial expression (Tye et al., 2014), suggesting that positive emotions are processed faster than negative emotions, for both groups. De Jong et al. (2008), Wagner et al. (2013) and Dawson et al. (2004) found that, unlike ASD individuals, TD individuals show differential responses to fear and neutral faces. Jamal et al. (2014) also found differences in emotion processing between groups, allowing group discrimination using Support Vector Machine classifiers. Akechi et al, (2010) and De Jong and colleagues (2008) demonstrated that the combination of facial expression and gaze direction modulates neural activity only in the TD group. Rapid processing of emotion is believed to rely on low spatial frequencies, which carry global stimulus information. However, subjects with ASD showed no emotion effect in low spatial frequency condition, but in high spatial frequency condition (Hendrika et al., 2010).

All articles that reported differences between hemispheres showed that ERP components had greater amplitude and lower latency in the right hemisphere (Akechi et al., 2010; Apicella et al., 2013; Dawson et al., 2004; O'Connor et al., 2005; Tye et al., 2014; Wong et al., 2008). Gross et al. (2012) found a reduced gamma activation of parieto-occipital cortices in individuals with ASD, suggesting a disruption of neural signaling in areas responsible for emotional facial expressions recognition.

Lerner et al. (2013) found that latencies of ERP markers of face processing were related to behavioral measures such that faster responses were associated with somewhat improved performance. The same finding was achieved by Garman et al. (2016) and Dawson et al. (2004). O'Connor et al. (2005) reported that TD and AS children do not differ in their ability to recognize basic expressions, but adults do.

In summary, this review provides strong evidence of group differences between ASD and TD individuals during facial expression recognition tasks, and right hemisphere dominance. However, the evidence in this field is still scarce, with several studies reporting results not replicated in the other studies. There is a clear need for further investigation on the effects of age and the dynamics of facial expressions, since most of the studies used static images as stimuli.

3.2 Neurofeedback and Brain Computer Interfaces in Autism Spectrum Disorder

Friedrich et al. (2014) highlighted the importance of using innovative noninvasive approaches, like BCI and neurofeedback, to provide insight about the physiological correlates of ASD. As these individuals often show a strong motivation towards technologies, the use of these interventions can be particularly effective, when compared to behavior and psychopharmacological interventions (which can be associated with side effects and long-term treatment) (Coben et al., 2010; Wainer and Ingersoll, 2011).

Friedrich et al. (2015) implemented a neurofeedback training based on the modulation of the mu rhythm, an 8-12 Hz oscillation, recorded over somatosensory cortex. Changes in this rhythm amplitude are closely related to social interactions, which, as stated before, are aspects in which ASD individuals are impaired (Friedrich et al., 2015). The authors developed a video game, which included social interaction and non-social interaction sequences. In non-social interaction sequences, child's avatar gets invited by a game character and they start the game picking apples and collecting coins to carry on their journey. During social

interaction sequences the child's avatar approaches the non-player character and, while facing this character, the participant must modulate his mu power in order to get rewarded. ASD children were rewarded if they decrease the mu power during social interaction sequences and increasing it in the non-social episodes. The rewarding feedback during social interactions consisted on the imitation of the facial expressions of the game character by the child's avatar. Results suggested that the game was successful improving brainwave responses, behavior and emotional reactions during social interactions; suggesting that this is a powerful tool for intervention in ASD individuals.

Virtual Reality (VR)-based interventions have also been investigated for ASD individuals (Fan et al., 2015). This technology creates immersive, interactive and realistic safe learning environments for individualized treatment. Fan et al. (2015) developed a VR-based driving system with EEG data acquisition for the purpose of addressing engagement level, emotional states and mental workload while driving. A future goal of the study was to incorporate an EEG-based BCI into the VR system to improve the system efficiency through individualized system adaptation based on the EEG signal for driving skill learning.

3.3 Electrophysiological Biomarkers in Autism Spectrum Disorder

In 2001, the Biomarkers Definitions Working Group defined a biological marker, biomarker, as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention" (Atkinson et al., 2001). Thus, a biomarker can be used as a diagnostic tool, as a tool for staging a disease, as an indicator of the disease prognosis or to monitor the clinical response to an intervention (Atkinson et al., 2001). Currently, the diagnosis of ASD can only be determined by clinical judgment, and no definitive physiological biomarker for ASD exists (Matlis et al., 2015). Thus, an electrophysiological biomarker, deduced from scalp EEG, could be important to help clinicians in the diagnosis of ASD, since

EEG could assess early neurophysiological dysfunctions that may not be apparent at behavioral level (Matlis et al., 2015).

In recent years, there has been an increase in the number of studies with interest in the development of a diagnostic biomarker of ASD using resting-state EEG data (Bosl et al., 2011; Duffy and Als, 2012; Heunis et al., 2016; Matlis et al., 2015; Maxwell et al., 2015; Pistorius et al., 2013; Tierney et al., 2012; Jun Wang et al., 2013). Resting state EEG studies allow to extract diagnostic biomarkers which are task independent, allowing to study the abnormal maturational trajectory in ASD through early childhood (Jun Wang et al., 2013). However, the existing studies have mixed results and lack of validation procedures, which limit their interpretation. Only few studies used statistical techniques and machine learning methods to validate their potential biomarkers of ASD diagnosis identified in the exploratory analysis (Duffy and Als, 2012; Matlis et al., 2015; Pistorius et al., 2013) and as a biomarker of ASD risk (Bosl et al., 2011).

Matlis et al. (2015) used spectral and functional connectivity methods to analyze EEG data from children with ASD and TD matched controls in an exploratory analysis. From the exploratory analysis, the authors hypothesized three possible biomarkers of ASD diagnosis: reduced posterior/anterior power ratio in the alpha frequency band in ASD individuals, relative to TD; reduced functional network density in ASD; and reduced mean connectivity strength in ASD group, relative to a subset of spatial locations (a subset of locations of the significant differences between ASD and TD networks). These biomarkers were tested in a validation group and the first and the third biomarker were validated, providing a significant classification between TD and ASD individuals.

Apart from the traditional EEG methods of analysis, EEG biomarkers can include complex signal processing and mathematical methods in time series analysis that are able to account for the nonstationary, nonlinear and complex nature of the EEG signals (Heunis et al., 2016). Heunis et al. (2016) reviewed three methodological approaches to resting-state EEG biomarker development for early detection of ASD: multiscale entropy, coherence analysis and recurrence quantification analysis.

Multiscale entropy, as a measure of signal complexity, was investigated by Bosl et al. (2011) in infants at high risk for ASD versus TD infants. Multiscale entropy quantifies the unpredictability of a time series across several time scales

(Jeste et al., 2015). The authors found multiscale entropy to be decreased in infants at high risk for ASD. Using multiscale entropy as feature vector, the authors used a support vector machine algorithm to classify between the two groups for different age groups, from 6 to 24 months. Differences between the two groups appear to be greatest at ages of 9 to 12 months, suggesting complexity measure goes through different developmental trajectory in infants at high risk for ASD and TD controls.

Duffy and Als (2012) investigated functional connectivity and found stable coherence patterns that significantly separate ASD and TD groups by discriminant function analysis. Coherence provides a measure of the degree of synchronization between two EEG signals (Heunis et al., 2016). Moreover, discriminant functions reliably classify individual control- and ASD-group subjects using leave-one-out technique. The authors found reduced short-distance and both reduced and increased long-distance coherences in ASD subjects, when compared with TD.

Pistorius et al. (2013) implemented recurrence quantification analysis, which is a measure of the system's complexity. This feature yielded information related to the complexity associated with the neural dynamics of each individual, enabling prediction of membership of each subject to ASD and TD groups.

The study of Bosl et al. (2011) investigated infants at high-risk of ASD, rather than children with a confirmed diagnosis of ASD. Studies like this (Tierney et al., 2012) aimed to investigate precursors of autism symptoms in infancy and hoped that valid biomarkers that are identified before the onset of clear symptoms will help in the clear detection of emerging autism (P. Walsh et al., 2011). Although these findings were proposed as an early diagnostic biomarker, and prevalence of ASD in siblings of a child with ASD being around 10 % (against the 1 % of the prevalence of ASD in general population), the large majority of these subjects does not develop ASD. Thus, the results of those studies need to be evaluated in follow-up investigations that include the subjects who developed ASD in order to assess a clear ASD biomarker.

Chapter 4

Methods

4.1 Participants

Seventeen male teenagers with ASD, recruited from the Unit of Neurodevelopment and Autism from the Pediatric Hospital of Coimbra and from the local Portuguese Association for Developmental Disorders and Autism of Coimbra and Viseu, participated in this study. The diagnosis was confirmed by DSM-5 criteria and Autism Diagnostic Observation Schedule (ADOS) and/or Autism Diagnostic Interview-Revised (ADI-R). Seventeen male TD teenagers also participated in this study, recruited from the local volunteers' database.

Participants from both groups were psychologically evaluated to assess their full scale and performance IQ. Cognitive ability was assessed by the Wechsler Adult Intelligence Scale (WAIS) for participants older than 15 years old, and by the Wechsler Intelligence Scale for Children (WISC) for younger participants. The two groups were matched by age and performance IQ.

Table 4.1 – Group characterization: Mean (standard deviation) of Age, Full Scale IQ (FSIQ) and Performance IQ (PIQ).

	N	**Age**	**FSIQ**	**PIQ**
ASD	17	16.4 (2.5)	92.2 (12.7)	99.8 (12.3)
TD	17	15.5 (2.6)	107.9 (18.0)	105.9 (18.0)

4.2 Stimulus and Task

Assessing facial expression recognition impairments in ASD using static photographs of facial expression stimuli, as most studies in this area do, may represent a problem, because this type of stimulus could be learned explicitly in the context of therapeutic interventions. In daily life, people encounter subtle facial expressions more often than those presented in static photographic stimuli. Furthermore, even when individuals with ASD demonstrate intact performance on facial expression recognition tasks using static photographic stimuli, as some of the studies report, they are likely to have problems in everyday social interactions. Thus, morphing facial expressions stimuli, that represent expressions of differing intensity (like morphing from neutral to facial expressions), are more likely to induce the same response as subtle facial expressions in daily social interactions and, because of that, may be more sensitive to differences between ASD and TD individuals (Harms et al., 2010).

In this study, subjects performed a visual task divided in two parts. A first part with visual stimulation, and a second with mental imagery of a third person performing a facial expression, like in visual stimulation. The full length of the experiment was around 30 minutes. Both tasks used slow dynamic morphing of form neutral to happy or sad facial expressions as visual stimulus. The visual stimulation and overall experiment were developed in WorldViz® *Vizard VR Toolkit* (development edition) using the *male002* virtual avatar from the *Complete Characters HD pack* and its facial expression poses.

The visual stimulation part consisted in visually present a virtual avatar dynamically morphing happy or sad facial expressions to the participants. Figure 4.1 presents the structure of the visual stimulation paradigm. Each stimulus has 1.5 s duration, and is always preceded by a baseline time of 1 s and a jitter of 500 ms, to avoid preparation. This part of the experiment consisted in two blocks of 120 randomized trials (60 of each condition).

Figure 4.1 – Schematic representation of the visual stimulation paradigm. In this part the subject had to look at the screen while dynamical morphing of happy and sad facial expressions is presented randomly. Each stimulus has 1.5 s duration, composed by a morphing period of 250 ms, a static period where the virtual avatar is displaying the facial expression for 1 s and a final period where the avatar morphs back to the neutral expression, with the duration of 250 ms.

The mental imagery part used, as instruction, a facial expression stimulus similar to the visual stimulation part. An auditory *beep* triggers the beginning of the imagery process, which takes four seconds, in which subjects had to imagine mentally the avatar performing the facial expression presented in the instruction. After the mental imagery phase, another *beep* is played to start a neutral phase with no imagery, which lasts also for four seconds. In this last phase, the subject had to stop the imagery period and to just look at the screen while waits for a new visual instruction. Figure 4.2 presents the structure of the mental imagery paradigm. The computer screen shows the neutral face of the avatar during the all period, except for the instruction. This task consisted in two blocks of 40 randomized trials (20 for each expression), for a total of 80 trials for the task.

As said, each subject performed two runs from visual stimulation and mental imagery parts. Between runs participants rested for a few minutes to ensure focus and reduce fatigue throughout the experiment.

Figure 4.2 – Schematic representation of the mental imagery paradigm. First, the dynamic morphing of happy or sad facial expressions similar to the visual stimulation part appears as instruction to the imagery process, lasting for 1.5 s. After the visual instruction, an interval of 1.5 s is left for preparation, and an auditory *beep* indicates the start of the mental imagery process, which takes 4 s. After that, a lower pitched *beep* is played to indicate the end of the mental imagery and the start the neutral period, with no imagery.

4.3 Experimental Setup and Data Recording

The visual stimulation was conducted in a 22-inch LCD Monitor (frame rate of 60 Hz, 1680x1050 resolution). The participant sat about 60 cm away from the screen (distance measured from the eyes to the center of the screen) and was asked to keep his eyes open and fixed on the face of the avatar, to reduce eye movement contamination. We recorded EEG data using a Brain Products® Package.

For data acquisition with an EEG system, firstly, individuals' scalp had to be exfoliated with an abrasive gel. Secondly, scalp and skin areas in contact with electrodes had to be cleaned with alcohol to ensure low impedances between the two surfaces in contact. We then positioned the Brain Products® actiCAP with 64 electrodes in subjects' heads, ensuring that the midline row of electrodes was properly aligned on the head and that electrode Cz was placed in head's vertex. Figure 4.3 illustrates the localization of the electrodes in actiCAP. This cap supports active electrodes based on high-quality Ag/AgCl sensors with integrated noise subtraction circuits delivering low noise levels. To create a conductive medium between scalp ad electrodes, we loaded all the electrodes with electrode

gel (see Figure 4.4). This gel allows to reduce the impedance of the brain signal to a value acceptable for the acquisition. The system measures each electrode's impedance and displays it at each electrode by an LED; when the value of the impedance is acceptable, the LED turns green. The electrode impedances for all channels were kept below 10 KΩ during the recordings.

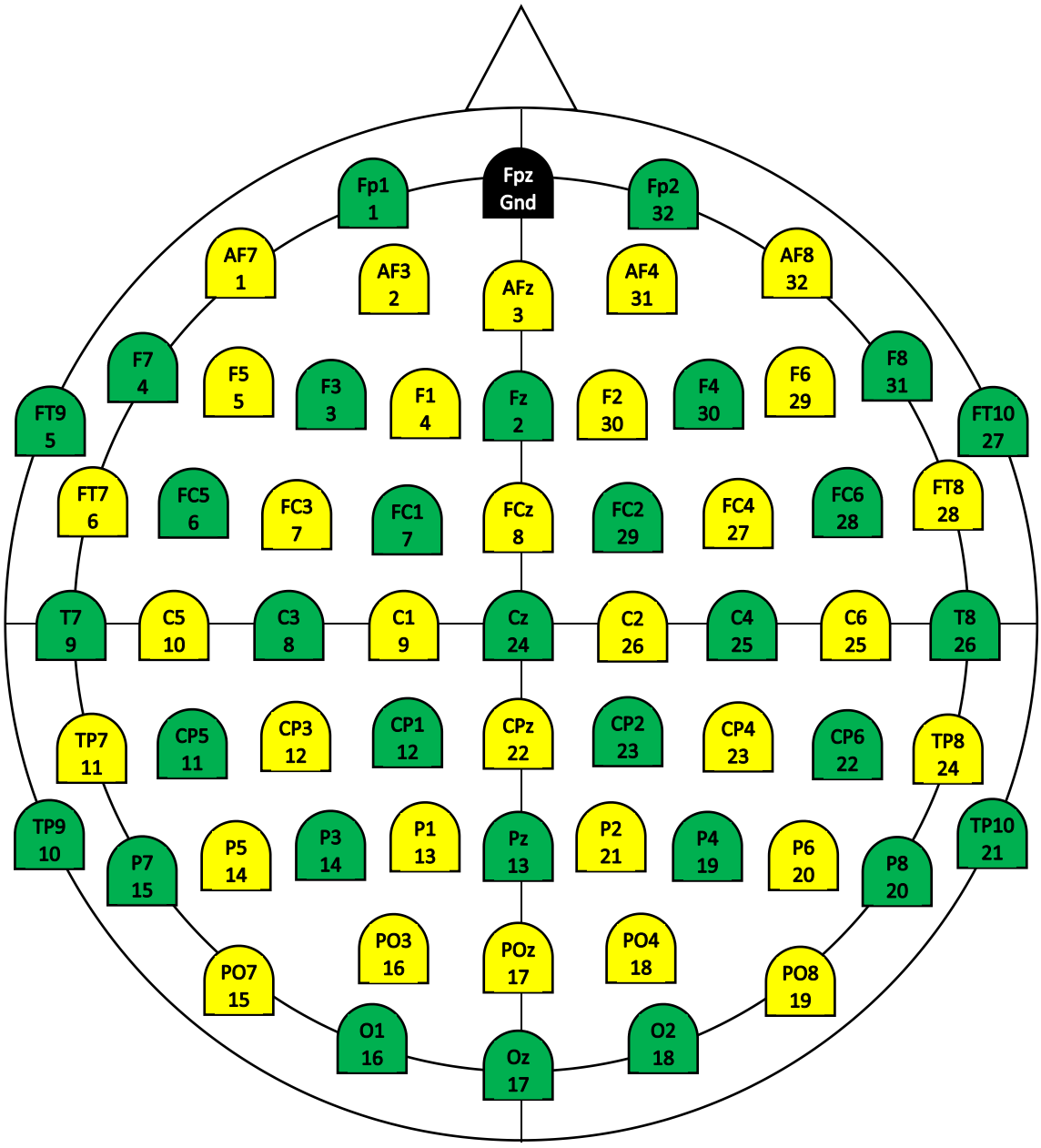

Figure 4.3 – Illustration of the location of each one of the 64 possible electrodes in actiCAP. Adapted from (Brain Products GmbH, 2011).

Figure 4.4 – Representation of Brain Products ® actiCAP with 64 possible active electrodes, from BrainProducts (Brain Products GmbH, 2016).

We recorded the EEG data using fifty-eight active electrodes, according to the expanded version of the International 10-20 system, the 10-10 system. We chose to place the ground electrode at AFz position and the reference electrode at the left earlobe. This electrode is usually placed in a location with minimal cortical activity and is used as a reference to subtract out correlated sources of noise. A set of four electrode positions recorded vertical and horizontal eye movements (electrooculography – EOG). The signal from those electrodes was posteriorly removed during the preprocessing procedures.

We acquired signal at a sampling rate of 1000 Hz using Brain Products® actiCHamp Amplifier. Brain Products® Brain Recorder software recorded the EEG data while the stimuli were presented to the subjects. The software used for stimulation generated different trigger pulses at the onset of each trial, for every instant a stimulus was presented. The trigger pulses were sent for the acquisition software to allow further processing procedures.

Before each recording, we instructed the participants to avoid talking, blinking or do any facial movements during the recording times due to bioelectrical artifacts that could compromise the acquired data. Each session, accounting for preparation and recording time took about 70 minutes.

4.4 Data Pre-processing

After storage of the raw data, we performed further processing procedures off-line, using MathWorks® Matlab R2013a and the EEGLAB toolbox v13.3.4b. We pre-processed EEG data and cleaned it from noise and artifacts. We applied the pre-processing chain, illustrated in Figure 4.5, equally in visual stimulation and mental imagery data.

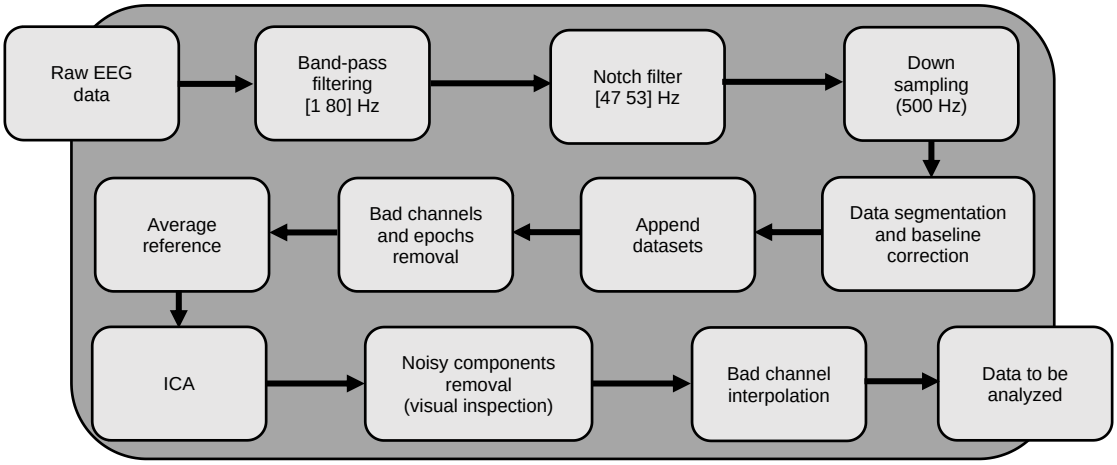

Figure 4.5 – Schematic illustration of the pipeline for the pre-processing of the EEG signals, from the acquisition data until it is cleaned of noisy artifacts.

Firstly, we filtered the data with a high and a low pass filter, to remove all the frequency components below 1 Hz and above 80 Hz, respectively. Additionally, we applied to the data a notch filter with a stopband of 47 to 53 Hz. This filter sharply attenuates electrical noise that occurs around 50 Hz. Before any other pre-processing procedure, we down sampled EEG data to 500 Hz.

Next, we segmented the data into epochs using the trigger information of each event. For happy and sad events in the visual stimulation task, we performed the segmentation in epochs of 1350 ms, from 100 ms pre-stimulus to 1250 ms post-stimulus. We also applied a baseline correction procedure on visual stimulation data, by subtracting the mean of the 100 ms pre-stimulus interval from the data after the stimulus onset. For the mental imagery task, we performed the

segmentation in epochs of 12000 ms with a 3500 milliseconds pre-stimulus interval, comprising baseline, visual instruction and preparation time, and 8500 ms post-stimulus interval, comprising mental imagery and no-imagery/resting times. Then, we appended in a single dataset the two EEG datasets representing each run, for visual stimulation and mental imagery tasks.

Data from EOG channels, that measured signals from eye movements were excluded from further analysis. Additionally, for cleaning noise and artifacts, we removed bad epochs and particularly noisy channels through visual inspection of the data. For bad epochs we considered those that had considerably large periods of noise or artifacts which might compromise further data analysis.

After removing noisy channels, we re-referenced the data with an average reference, in which an average of all the recordings on every electrode site was taken and used as a reference. Through this process, only signal/noise that was common to all sites remain (correlated); the contribution uncorrelated noise sources were minimized (Ludwig et al., 2009).

We used the Independent Component Analysis (ICA) decomposition method to linearly transform the multi-channel EEG data recorded into a collection of components that are statistically independent from each other (Erfanian et al., 2011). These independent components correspond to the outputs of spatial filters applied to the whole multi-channel data, which is constituted, in fact, by many sources that had been mixed via volume conduction and had been recorded at scalp channels. These signal sources may represent synchronous activity within one or more cortical patches, or activity from non-cortical sources, constituting artifacts. These non-cortical sources could be eye movements or eye blinking, resulting in eye artifacts that appear in multiple channels; teeth clenching and muscle tension; pulse signal; or single channel artifacts (when a single channel goes off only contributing with noise). The components corresponding to these types of artifacts are in frequency and voltage range (amplitude) of EEG signals, being detected by scalp electrodes. Thus, they should and were removed in order to clean the EEG signal. Also, through visual inspection of the EEG signals and ICA components, we removed additional components associated with some kind of noisy pattern that appeared in the channel data.

We interpolated signals from removed electrodes using a spherical method. Thereafter, we split the dataset corresponding to visual stimulation task in two

different files by selecting the specific trials' indexes corresponding to each condition defined by the triggers (happy or sad). We applied the same procedure to the mental imagery task.

4.5 Data Analysis

In the analysis of the mental imagery task, we used a linear Support Vector Machine (SVM) to discriminate:

- No-imagery states from facial expression imagery states,
- ASD from TD individuals, based on the data of the mental imagery period.

These two different analysis used different processing and feature extraction methods. As such, we performed different feature selection procedures and classification analysis in the two different datasets. Thus, the processes will be described independently in two sections.

4.5.1 Classification of Mental Imagery of Facial Expressions *versus* No-Imagery

For the discrimination between mental imagery and neutral/no-imagery states, we extracted features from these two different time intervals for each trial of each subject. Figure 4.6 summarizes the key steps taken in this classification system.

We used a 5-fold cross validation technique, and applied a feature extraction procedure on the number of trials of each subject dataset, which were afterwards used to estimate the accuracies of each subject's classification. This implies that for each subject, we performed feature extraction procedure five times, using different sets of trials' indices to build train and test data sets. At the end, there

were five different train and test sets with as many observations per condition as the number of trials selected for train and test, respectively.

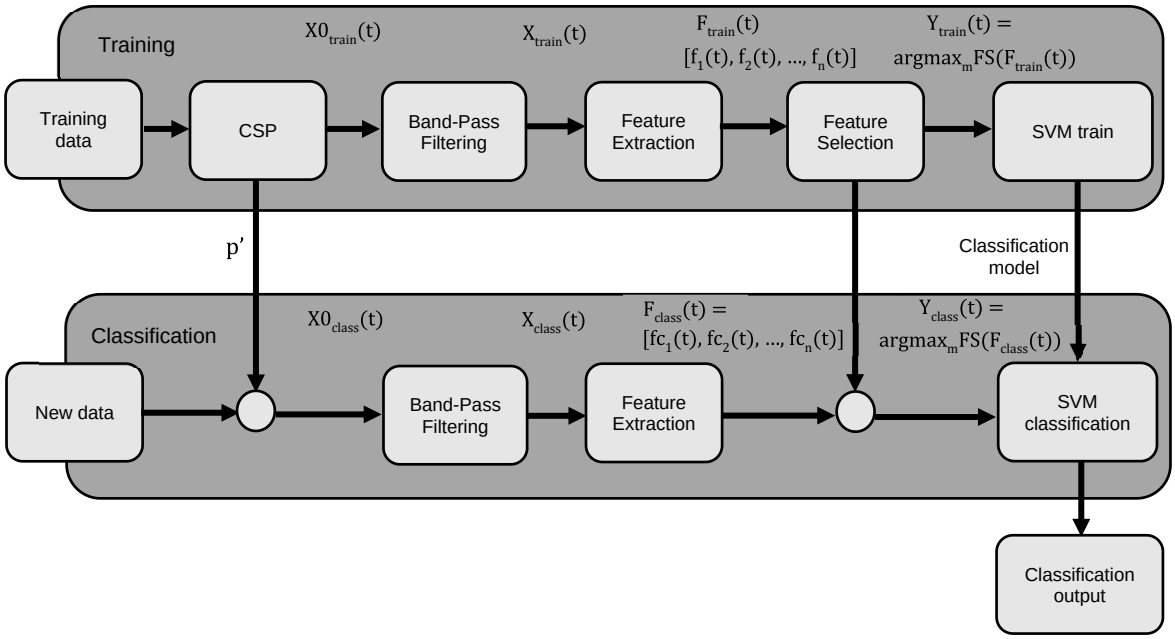

Figure 4.6 – Schematic illustration of the pipeline for this classification system. Firstly, we divided EEG data into training and test datasets, through a 5-fold cross validation technique. After the application of the CSP algorithm to the training dataset we extracted the 5 most significant patterns. Then, we extracted the defined features from the band-pass filtered data. To select the best features, we applied a feature selection algorithm. The best features were to build a linear SVM model for this classification system.

We split the train datasets into four new subsets, corresponding to the time intervals of the two conditions tested and their respective baseline intervals. We performed the same procedure in test datasets. The epochs extracted for the two condition times were of the same size; the same applies to the size of the two baseline intervals. We tested different conditions as well as different time windows for baseline. For the comparison of mental imagery of facial expressions with no-imagery states we used the time intervals of [500 4000] ms and [4500 8000] ms, respectively; the baseline intervals were the same for both conditions, but we tested these two hypothesis: [-3500 -3000] ms (before visual instruction period) and [-450 0] ms (preparation period). Initially we chose to perform the tests using

the baseline interval during the preparation period, but, since the results were not very good, we discussed if the individuals could had started the mental imagery task just after visual instruction, and suggested a new baseline interval, before the visual instruction period.

To reduce dimensionality of the data to use in feature extraction, we applied the Common Spatial Patterns (CSP) algorithm. CSP is a mathematical procedure commonly used in electroencephalographic data processing for separating multichannel data into additive subcomponents, which have maximum differences in variance between two time windows. We applied CSP algorithm to the train dataset in order to maximize the differences in variance between the signal of the two different epochs. Mathematically, let X_1 of size (n, t_1) and X_2 of size (n, t_2) be two windows of a multi-channel data, where n is the number of signals (number of channels), and t_1 and t_2 the respective number of samples (time points of the windows extracted previously). The CSP algorithm determines the component w^T that maximizes the ratio of variance between the two windows, given by,

$$w = \arg max_w \frac{\|wX_1\|^2}{\|wX_2\|^2} \tag{4.1}$$

The solution is given by computing the two covariance matrices:

$$R_i = \frac{X_i X_i^T}{trace(X_i X_i^T)}, i = 1, 2 \tag{4.2}$$

The elements of R thus represent a measure of the fractional variance of each EEG channel in an epoch and the off-diagonal elements represent the fractional covariance (Koles et al., 1990). From the covariance matrices R_i for each selected epoch, a population of covariance matrices R_p is computed as:

$$R_p = \frac{1}{N_p} \sum_{i=1}^{N_p} R_i , \tag{4.3}$$

where N_p is the number of EEG epochs in the population p (Koles et al., 1990). Then, the eigenvalue decomposition is performed and the eigenvectors $P = [p_1, \dots, p_n]$ are determined, and the diagonal matrix of eigenvalues $D = \{\lambda_1, \dots, \lambda_n\}$ sorted by decreasing order, such that:

$$R_p = P_p D_p P_p^T, \qquad P_p^T P_p = I \tag{4.4}$$

The eigenvectors composing P are components with variance ratio between the two windows equal to their corresponding eigenvalue. So, the component with the maximum ratio of variance is given by the first column of P, $w = p_1^T$, as it corresponds to the first eigenvalue, and so on, since the diagonal matrix of eigenvalues are sorted by decreasing order.

We, then, reduced the size of the train dataset from fifty-eight channels to the five most significant patterns extracted from CSP. We also reduced the size of baseline and test datasets using the corresponding data of the five most significant patterns selected.

After the application of CSP to the data and selection of the significant patterns to use, we filtered the resultant EEG data from the two windows and their respective baselines in seven different frequency bands using a band-pass filter. The frequency bands selected for analysis were: theta ([4 8] Hz), alpha ([8 15] Hz), beta ([15 30] Hz), beta sub-bands ([15 20] Hz, [20 25] Hz and [25 30] Hz), and gamma ([30 40] Hz). Thereafter, EEG data were ready for feature extraction.

4.5.1.1 Feature Extraction

We selected 35 different features for extraction from the source signal segments. Those features were grouped in two categories, time-domain features and non-linear features. In each dataset, we calculated each feature for each trial, frequency band and CSP component.

Frequency domain information is the most widely used method in clinical applications of BCI systems with neurofeedback. However, linear features cannot represent the brain activities due to the nonlinearity of the EEG signal itself. While mental tasks are being performed, EEG segments of any length started from a

specific time point are related to different kinds of brain activity information. Non-linear methods, like entropy analysis and phase space characteristics became popular in many EEG processing for medical applications and could be applied to neurofeedback systems to model brain activities (Q. Wang et al., 2010).

Simões et al. (2015) previously validated the features extracted, as they were used in the same task but in an EEG-fMRI environment.

4.5.1.1.1 Time-Domain Features

- Hilbert Envelope (Env) – extracts the envelope (smooth curve outlining the extremes of the signal), which corresponds to the magnitude of the analytic signal. The analytic signal is composed by the original waveform and its Hilbert transformation. Hilbert transformation of the signal corresponds to the original waveform with a 90º phase shift. Mathematically, the analytic signal $z(t)$ can be defined by $z(t) = x(t) + i\hat{x}(t)$, in which $\hat{x}(t)$ corresponds to the Hilbert transformation of the original signal $x(t)$ (Simões et al., 2015; Ulrich, 2006);
- Power (Pow) – the sum of the absolute squares of time-domain samples of the signal, divided by the signal length (Mathworks, 2016; Simões et al., 2015). Mathematically, the power P of a signal $x(t)$ over all time is defined by

$$P = \lim_{N\to\infty} \frac{1}{2N} \sum_{n=-N}^{N-1} |x[n]|^2 \quad (4.5)$$

- Teager Energy (Teag) – an energy estimation operator which uses the sum of the instantaneous energy of the signal divided by the signal length. The instantaneous energy of a signal x_n can be determined using $\psi[x_n] = x_n^2 - x_{n-1}x_{n+1}$, (Antoniadou et al., 2012; Simões et al., 2015).

For each one of these features, we extracted maximum, minimum, average and standard deviation values. Additionally, we extracted these feature values from baseline segments. The final vector of time-domain features included absolute and

baseline-corrected values of maximum, minimum, average and standard deviation of Env, Pow and Teag. We performed baseline correction by subtracting, to the absolute value of the feature, the value of the feature during the baseline segment

4.5.1.1.2 Non-Linear Features

To extract signal complexity measures, we transformed the signal to phase space. Phase space allows to demonstrate and visualize the changes in the dynamical variables of the system. Every possible state of the system can be represented by a point in the multidimensional phase space and time evolution of the system creates a trajectory in the phase space (Kliková and Raidl, 2011). Figure 4.7 represents an example of a trajectory of a given signal in the phase space.

The method used to reconstruct the phase space of the signal was based on the method of time delay. Mathematically, given a time series of a scalar variable it is possible to construct a vector $x(t_i),\ i = 1, \dots, N$ in phase space in time t_i as following:

$$X(t_i) = [x(t_i), x(t_i + \tau), \dots, x(t_i + (m-1)\tau)], \quad i = 1, \dots, N - (m-1)\tau \tag{4.6}$$

where τ is time delay, m is a dimension of reconstructed space and $M = N - (m - 1)\tau$ is the number of points (states) in the phase space.

It was reconstructed a 2 and 3-dimentional phase space associated to the EEG data, and we considered the time delay to be mean of the first local minimum from the signal's autocorrelation (Lag).

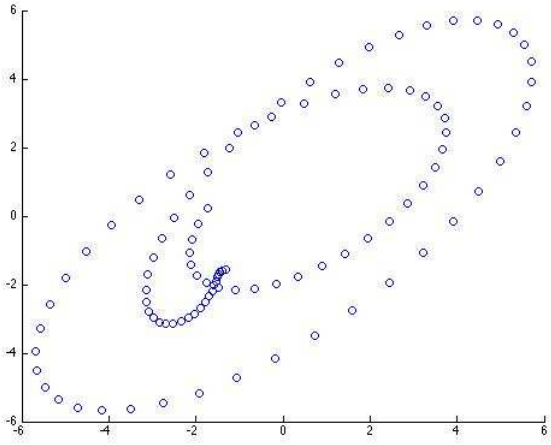

Figure 4.7 – Example of a trajectory of a signal in the phase space.

The reconstruction of the phase space allows to estimate non-linear measures like:

- Largest Lyapunov Exponent (Lyap) – characterizes the rate of separation of infinitesimally close trajectories of the signal in phase space, providing a measure of the degree of the system's instability (Cencini, Cecconi, and Vulpiani, 2010). Mathematically, two trajectories in the phase space with initial separation of δZ_0, diverge at a rate given by

$$|\delta Z(t)| \approx e^{\lambda t}|\delta Z_0| \tag{4.7}$$

  where λ is the local Lyapunov exponent (local exponential rate of expansion) (Cencini et al., 2010). The rate of separation can be different for different orientations of initial separation vector, leading to a spectrum of Lyapunov exponents – equal in number to the dimensionality of the phase space. The maximum Lyapunov exponent corresponds to the mean exponential rate of divergence, characterizing the trajectory's instability (positive values are associated with a chaotic system) (Cencini et al., 2010);

- Correlation Dimension (CorrDim) – measure of the space dimensionality of the signal in phase space. Correlation sum is defined as sum the fraction of pairs of points of the phase space whose distance is smaller than r, being r the Lag (defined previously). If this number of points is sufficiently large, the ratio between the logarithm of the correlation sum and logarithm of the time delay is a good estimate of the Correlation Dimension (Cencini et al., 2010).
- Approximate Entropy (ApEn) – quantifies the amount of the regularity and unpredictability of fluctuations of the signal (Pincus et al., 1991). A time series with many repetitive patterns has a small value of ApEn, reflecting its predictability; the opposite happens for less predictable signals;
- Sample Entropy (SpEn) – is a modification of ApEn and is used for assessing the complexity of a physiological time series data. ApEn depends on the length of the time series and lacks relative consistency. SpEn, similarly to ApEn, quantifies the regularity of the signal but does not have the disadvantages mentioned (Richman and Moorman, 2000);
- Spatial Filling Index (SFI) – normalizing the phase space, the positions spans from -1 to 1 on either axis. The phase space area is then divided into small square areas of size $R \times R$, and the number of voxels in the normalized phase space is $n = \left(\frac{2}{R}\right)^m$, being m the number of dimensions of the phase space. We can obtain a new matrix with its elements equal to the number of phase space points falling in each grid (Faust et al., 2004). SFI corresponds to the probability of a phase space point falling in a grid.

Additionally to absolute value, we extracted difference to baseline values of these features and included in the final feature vector.

After feature extraction procedure, we obtained, for each subject and for each fold, 1225 features for each observation in train and test sets. As previously said, the amount of observations for each condition was the same.

In order to reduce the number of features for classification, and before applying a feature selection algorithm, we performed an average through the features of the five most significant patterns of CSP, for each feature and frequency band. The number of features after this procedure was reduced to 245, which corresponds to $35\ features\ \times\ 7\ frequency\ bands$.

After that, we performed a moving average through observations (trials) using a window of one, two and three trials. We performed this procedure for each one of the five folds, for each subject. A moving average consists in a calculation to analyze data points by creating a series of averages of different subsets of the full dataset, smoothing the data. For training sets, given a series of observations, the first element of the moving average is obtained by taking the average of the initial fixed subset of the observations (one, two or three trials); then, the subset to be averaged is modified, excluding the first number of the series and including the following number from the original subset in the series. This process is repeated over the entire length of the observations. For testing sets, the shifting performed to the subset to be averaged was different; after the first average, the next subset to be averaged does not contain any element of the previous subset, which reduces the number of the final observations depending on the window of samples to be averaged.

After performing the moving average procedure, the final number of train and test datasets increased to 15 for each subject, which corresponds to the three moving average procedures performed for each one of the 5-fold cross validation sets.

4.5.1.2 Feature Selection

Feature selection reduces the dimensionality of the data by selecting a subset of features from the input feature set. Usually, feature selection methods are used to reduce the computational cost and to remove the irrelevant and redundant features from the high dimensional input feature set (Gu et al., 2012).

Using the training sets of the moving average with one trial, we performed a feature selection procedure using Fisher Score algorithm to each one of the 5-fold datasets.

4.5.1.2.1 Fisher Score Algorithm

Fisher Score (FS) algorithm is one of the most widely used supervised feature selection methods. We used this algorithm to rank the features by relevance and the best features were then fed into the SVM classifier.

FS is a filter-based method which does a binary selection of the features according to some performance criterion. This algorithm ranks the features as a pre-processing step prior to the learning algorithm, and select those features with high ranking score (Gu et al., 2012).

Mathematically, given a set of d features, denoted by S, the goal of the filter-based feature selection is to choose a subset of $m < d$ features, denoted by Γ, which maximizes some performance criterion F,

$$\Gamma^* = \arg max_{\Gamma \subseteq S} F(\Gamma), s.t. |\Gamma| = m \quad (4.8)$$

where $|.|$ is the cardinality of a set (Gu et al., 2012).

Finding the optimal solution for this equation is hard. One common heuristic approach is to first compute the score for each feature independently according to the criterion F, and then select the top m ranked features (with higher scores) (Gu et al., 2012). However, this heuristic has some problems: it neglects the combination of features, since it scores them individually, and it cannot handle redundant features, attributing high score to a set of features even if they are all highly correlated (Gu et al., 2012).

For the implementation of FS for feature selection we used the method developed by Giorgio (2016). At the end, we averaged the ranking of the 5-folds training sets for each subject, obtaining the mean ranking for each subject. With this information, it was possible to perform the mean of the ranking across all the subjects and determine which were the best features in the ranking.

4.5.1.3 Support Vector Machines Classifier

For the classification procedure, we used a linear Support Vector Machines (SVM) classifier, built in the statistics toolbox of Matlab.

SVM is one of the most popular machine learning methods. This algorithm is often used for BCIs and to classify biological signals (Kashihara, 2014; Yoshimura and Itakura, 2011). Given a training set, with each observation labelled to one of two classes, an SVM training algorithm builds a model that allows for the classification of new testing examples, assigning each test observation into one class or the other. An SVM model consists on a representation of the observations of the training set into points in space, mapped so that the observations assigned to different classes are divided by a clear gap that is as wide as possible (Martínez and Barrientos, 2011). The support vectors are the points at the minimum distance from the hyperplane which separates the two classes. The SVM classification algorithm maps the test set observations into the same space and predicts which class they belong based on which side of the hyperplane they fall on.

Mathematically, in a linear SVM, given a training set X, with n observations and m variables, each x_i observation is a m– dimensional real vector. Each x_i has an attributed class y_i, which can be either 1 or -1. The objective is to find the maximum-margin hyperplane that divides the group of points x_i for which $y_i = 1$ from the group of points for which $y_i = -1$, defined in order to maximize the distance between the hyperplane and the nearest points (support vectors) from either group (see Figure 4.8).

Any hyperplane can be written as the set of points x satisfying $w.x + b = 0$, where w is the normal vector to the hyperplane. If the training data are linearly separable, is possible to select two parallel hyperplanes that separate the two classes of the data so that the distance between them is as large as possible. The maximum-margin hyperplane is the hyperplane that lies between the two parallel hyperplanes, defined by $w.x + b = 1$ and $w.x + b = -1$; the distance between these two hyperplanes is $\frac{2}{\|w\|}$. It is assumed that the bigger this distance is, the bigger the generalization capability of the model is (Martínez and Barrientos, 2011).

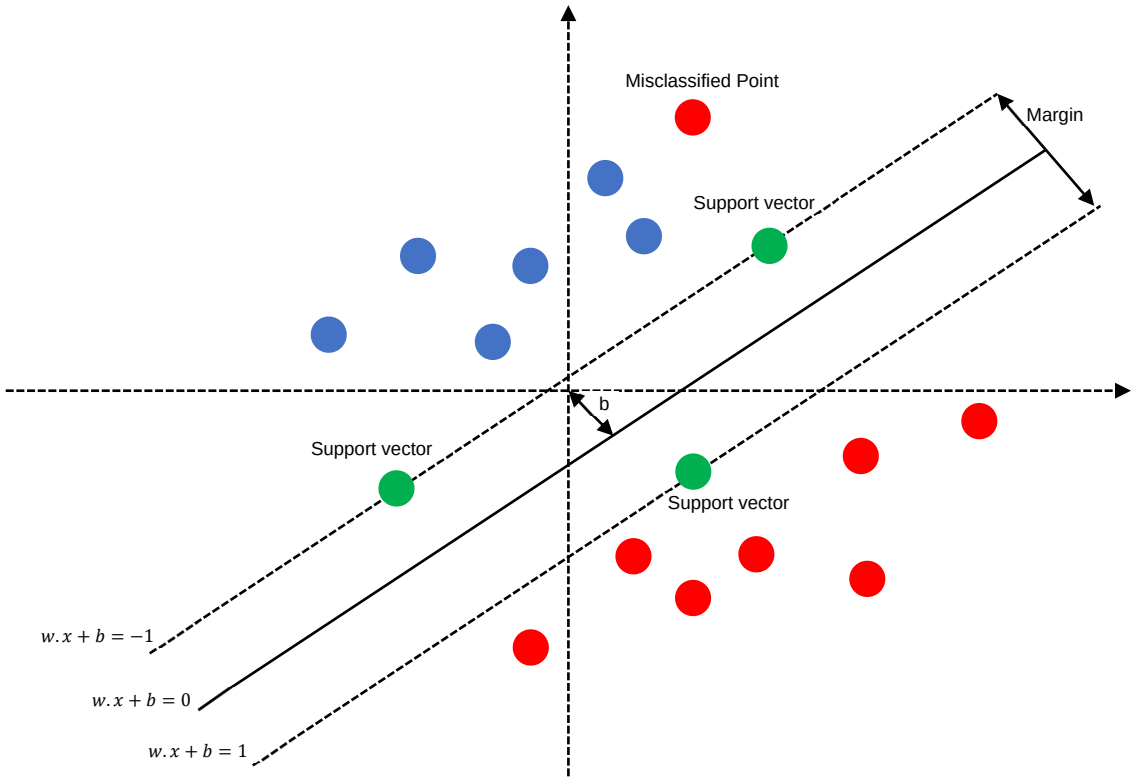

Figure 4.8 – Schematic representation of a linear SVM model. Blue and red dots correspond to data points from two different classes. Green dots represent the support vectors which define the maximum-margin hyperplane that separated the two classes.

In this case, for each fold, we used the training dataset to build the SVM model using the data corresponding to the n best features, with $n = 1, \dots, 245$. We determined the best features through averaging the ranking across all the 34 subjects. We then applied the SVM model to the test set (built with the corresponding data from the same best features). The original labels and the outputs of the classification for each one of the 5-fold sets were saved. We repeated this process to the data corresponding to each moving average procedure done (single trial, two and three trials).

For each window of the moving average (one, two and three), we determined the accuracy based on the results from the classification of the 5-fold sets. Also, we determined the p-value based on 10000 permutations for each window. We performed this procedure for each subject, which allowed to determine the average accuracy for each group of subjects.

4.5.2 Classification of ASD *versus* TD individuals

We performed this second analysis with the objective to find a biomarker or a set of biomarkers that may distinguish the two groups of subjects, using the mental imagery of facial expressions.

Figure 4.9 summarizes the key steps taken in this classification system.

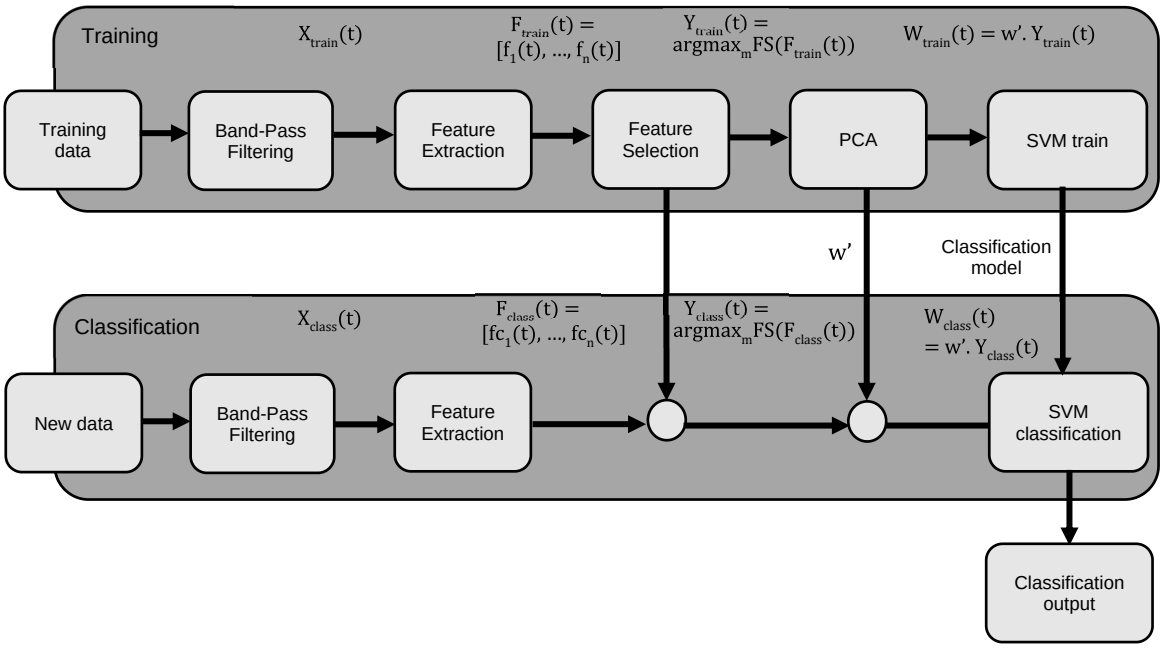

Figure 4.9 – Schematic representation of the pipeline for this classification system. Firstly, we filtered the EEG data into different frequency bands. Then, we extracted 35 different features from each data set. We applied a feature selection procedure followed by a PCA analysis in order to select the principal components used to build the linear SVM model for this classification system.

For each subject, from its original data, the time interval corresponding to the mental imagery state was extracted (500 to 4000 ms post-stimulus event), and its corresponding baseline (from 3500 to 3000 ms pre-stimulus event). We did the extraction of these epochs from a dataset containing both happy and sad condition trials and we did not make any differentiation between the two facial expressions.

We filtered the EEG data selected from the time intervals in seven different frequency bands using a band-pass filter. The frequency bands selected, like in the previous analysis, were: theta ([4 8] Hz), alpha ([8 15] Hz), beta ([15 30] Hz),

beta sub-bands ([15 20] Hz, [20 25] Hz and [25 30] Hz), and gamma ([30 40] Hz). Thereafter, EEG data were ready for feature extraction.

4.5.2.1 Feature Extraction

In this case, we extracted the same time-domain and non-linear features used in the previous analysis from the EEG filtered data. For each subject, we calculated each feature for each trial, frequency band and EEG channel. After feature extraction, we performed an average through the trials, obtaining only one set of features for each subject. At the end, each subject had a set of 14210 features that characterized him.

To reduce the dimension of the set of features, we defined a set of 13 channel clusters and we performed an average for each feature of each frequency band through the channels constituting each one of the 13 clusters. Figure 4.10 represents the locations of the 13 clusters defined for this analysis. With this average procedure we reduced the set of features for each subject to 3185 features, which corresponds to $35\ features \times 7\ frequency\ bands \times 13\ channel\ clusters$.

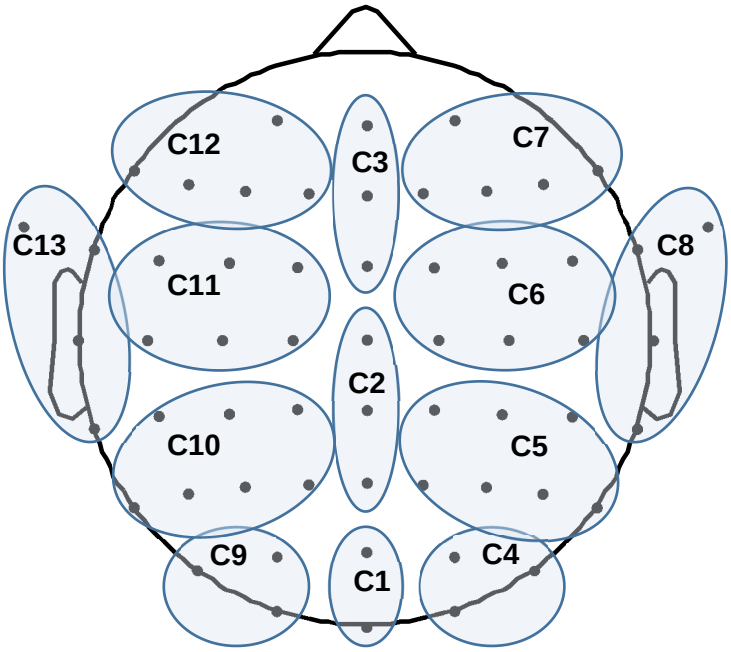

Figure 4.10 – Representation of the clusters selected for this analysis. Cluster 1: POz, Oz; Cluster 2: Cz, CPz, Pz; Cluster 3: AFz, Fz, FCz; Cluster 4: PO4, PO8, O2; Cluster 5: CP2, CP4, CP6, P2, P4, P6, P8; Cluster 6: FC2, FC4, FC6, C2, C4, C6; Cluster 7: AF4, F2, F4, F6, F8; Cluster 8: FT8, FT10, T8, TP8; Cluster 9: PO3, PO7, O1; Cluster 10: CP1, CP3, CP5, P1, P3, P5, P7; Cluster 11: FC1, FC3, FC5, C1, C3, C5; Cluster 12: AF3, F1, F3, F5, F7; and Cluster 13: FT7, FT9, T7, TP7.

4.5.2.2 Feature Selection

We applied Fisher Score algorithm to the data matrix $[34\ subjects \times 3185\ features]$ in order to obtain the ranking of the features extracted, and select the most relevant features to distinguish between ASD and TD individuals. We then preprocessed a data set with the best 100 features selected from the ranking in two different ways: Principal Component Analysis (PCA) and best cluster and frequency selection.

For the first approach, we applied a PCA to the data in order to reduce the dimensionality of the data and redundancy between features.

PCA is also a well-known feature selection method. Its purpose is to find an orthogonal set of projection vectors which are uncorrelated variables, called principal components, for feature selection from a given input set of observations of possibly correlated variables. This orthogonal transformation is defined in such a way through maximizing the variance of the projected data (the first principal component has the largest possible variance, that is, accounts for as much of the variability in the data as possible, and each succeeding component has the highest variance possible under the constraint that it be orthogonal to the preceding components). The resulting vectors are an uncorrelated orthogonal basis set, with aim of optimal representing the data in terms of minimal reconstruction error (Erfanian et al., 2011). If a multivariate dataset is visualized as a set of coordinates in a high-dimensional data space, PCA can supply a lower-dimensional view of the data using only the first few principal components, a projection of the data when viewed from its most informative viewpoint. However, this method does not use class information and so, its maximization of the variance of the projected patterns might not be necessarily a good contribute to the following of discrimination of classes. Thus, the projected data loses some useful discriminating information for classification (Erfanian et al., 2011).

Mathematically, given a data matrix X, with n – observations × m – variables, a set of k m–dimensional vectors of weights, $w_k = (w_1, \dots, w_m)_k$, $k \leq m$ (being k the number of principal components), can be defined in order to map each observation x_i, $i = 1, \dots, n$, of X, to a new vector of principal component scores $t_i = (t_1, \dots, t_k)_i$. The orthogonal linear transformation is then given by $t_{k(i)} = x_i . w_k$, in which the individual variables of t contains the maximum possible variance from x, with each loading vector w constrained to be a unit vector.

The second approach was to select from that set of 100 best features only the features correspondent to the best channel clusters and frequency bands. We selected the best channel clusters and frequency bands based on their distributions through the ranking. Then, to the new dataset built with the features from the best clusters and frequency bands, we applied a Principal Component Analysis to select the principal components which best explain the information present in the dataset.

4.5.2.3 Support Vector Machines Classifier

After obtaining the principal components of each dataset, we used a linear SVM classifier to distinguish ASD and TD individuals.

In this case, the classification procedure was based on a leave-one-out cross validation technique. For 34 times $(N = 34)$, we built a training dataset with $N - 1$ subjects (observations) and their x principal components, the subject left out was used as a testing set. We performed this procedure until all the subjects had been used as the test set. The original label of all the testing sets and the outputs of the classification were saved, in order to determine the accuracy of the classification. In addition, we determined the p-value associated to the accuracy result, based on 10000 permutations.

Chapter 5

Results

5.1 Classification of Mental Imagery of Facial Expressions

In this analysis, we explored different comparisons between conditions and different baseline intervals. Results presented in the next subsections correspond to the best accuracy values obtained and, because of that, they are based on one of the two baseline intervals tested: [-450 0] ms and [-3500 3000] ms, pre-stimulus. Mental imagery and no-imagery features were extracted from [500 4000] ms and [4500 8000] ms time intervals, respectively.

5.1.1 Mental Imagery of Facial Expressions *versus* No-Imagery

For this comparison, we extracted features from EEG datasets containing trials of both conditions (happy and sad facial expressions). The two baseline intervals were tested, and the best results of classification were obtained using as baseline interval time points from 3500 ms to 3000 ms pre-stimulus.

We applied Fisher Score algorithm to each subject's training data and obtained a feature ranking for each one. Additionally, we performed an average through the 34 subjects to determine the average ranking. We analyzed frequency bands distribution over average ranking for all features extracted, for time-domain features alone and for non-linear features alone. The box plots representing these

distributions are presented in Figure 5.1, Figure 5.2 and Figure 5.3. It is possible to note, in Figure 5.1, that all frequency bands have the similar median ranking positions. However, some frequency bands have narrower ranges of the ranking positions of their features than others.

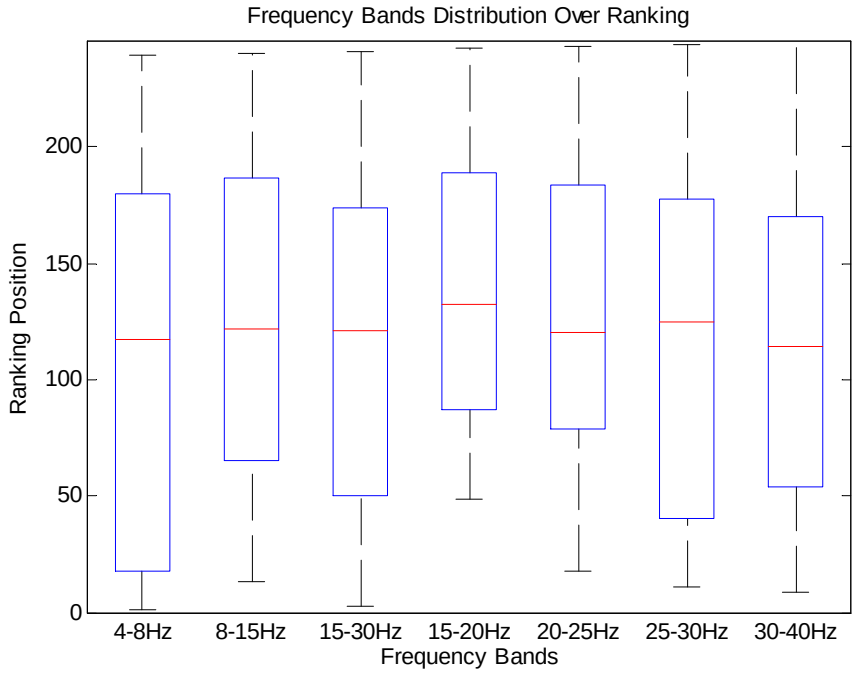

Figure 5.1 – Distribution of the different frequency bands over ranking. Ranking positions of the features corresponding to each frequency band were saved and plotted in a box plot graph. This box plot shows the median ranking position of each frequency band, as well as the range of their positions in ranking.

From the distribution of time-domain features, Figure 5.2, it becomes clear which are the frequency bands with the lowest mean positions in ranking. Until the 75th position theta and gamma appear to be the most common frequency bands. It is also clear the median ranking positions for that features appear to be close from 100, for all frequency bands, and, that 75% of the features (also for all frequency bands) are below 175th ranking position. In contrast, non-linear features (Figure 5.3) appear to occupy the highest positions in the ranking, not having much relevance to the correct classification of the observations.

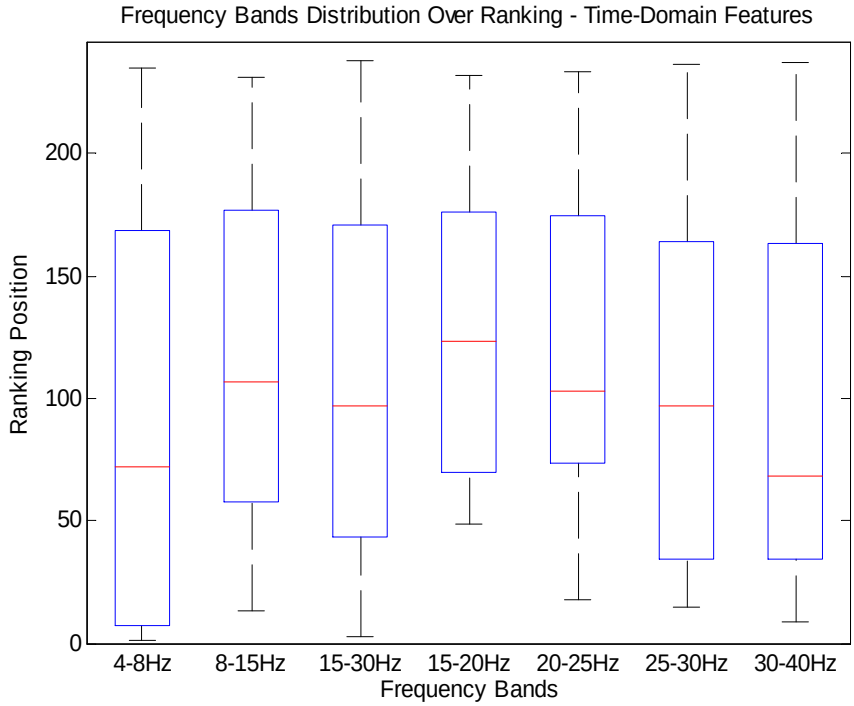

Figure 5.2 – Distribution of the different frequency bands over ranking. Ranking positions of the time-domain features corresponding to each frequency band were saved and plotted in a box plot graph. This box plot shows the median ranking position of each frequency band, as well as the range of their positions in ranking.

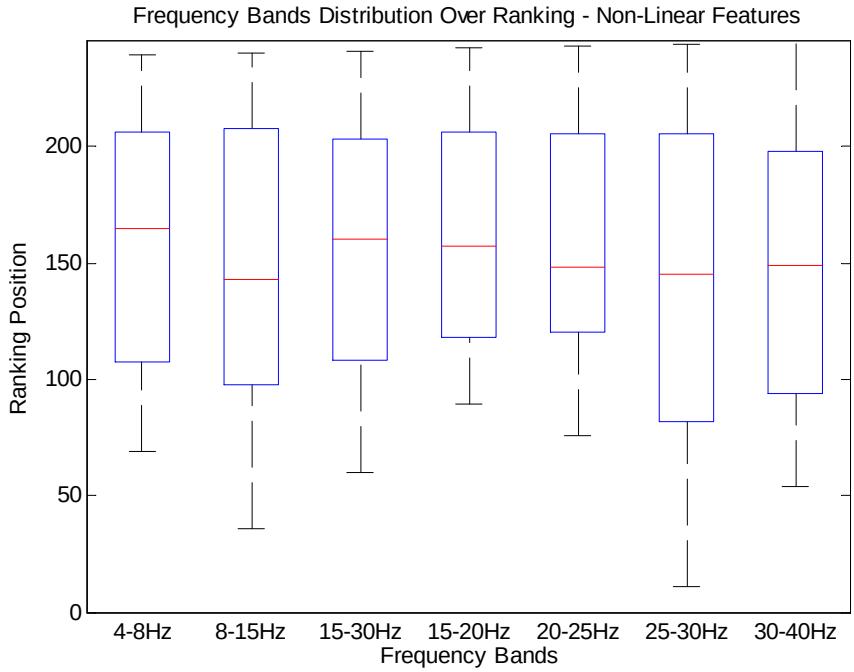

Figure 5.3 – Distribution of the different frequency bands over ranking. Ranking positions of the non-linear features corresponding to each frequency band were saved and plotted in a box plot graph. This box plot allows to know the distribution of ranking positions of each frequency band.

The type and frequency band of the best 10 features of the average ranking is presented in Table 5.1. All the 10 best features were from time-domain, as expected from the visual inspection of the information given from Figure 5.2 and Figure 5.3. Except for only one feature, the others were absolute values of the features extracted during the time intervals of both conditions. Most of the features present in this set were extracted from the theta band EEG signal, [4 8] Hz.

Table 5.1 – Characteristics of the 10 best features of the ranking.

Ranking Position	Feature				Frequency Band (Hz)
	Time-Domain		Non-Linear		
	Absolute	Baseline Correction	Absolute	Baseline Correction	
1	SD Env	-	-	-	[4 8]
2	SD Teag	-	-	-	[4 8]
3	Maximum Teag	-	-	-	[4 8]
4	-	Maximum Env	-	-	[15 30]
5	Average Teag	-	-	-	[4 8]
6	SD Pow	-	-	-	[4 8]
7	SD Pow	-	-	-	[30 40]
8	Average Pow	-	-	-	[4 8]
9	Average Env	-	-	-	[4 8]
10	SD Teag	-	-	-	[30 40]

We built a linear SVM model for the classification procedure, using the corresponding data from best features of the average ranking, and classification accuracies were determined through a 5-fold cross validation technique, as explained earlier (Section 4.5.1.3). We performed that procedure over the data resultant from the moving average of one, two and three trials, also as explained.

Figure 5.4 shows the accuracy progression using moving average data of one, two and three trials for both groups using from 5 to the 50 best features of the ranking. The accuracy value presented in the graph corresponds to the mean accuracy of the group, and the error bars to the Standard Error of the Mean (SEM), determined by

$$SEM = \frac{s}{\sqrt{N}} \tag{5.1}$$

in which $s = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$ corresponds to the Standard Deviation (SD) of the sample, and N to the size of the sample.

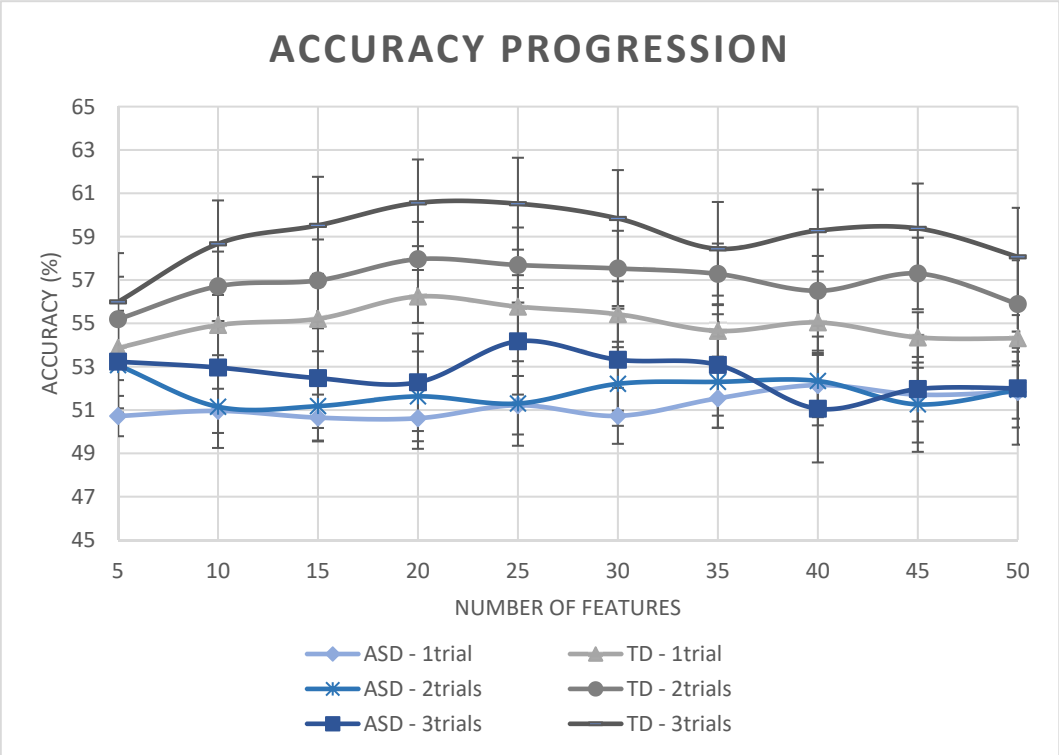

Figure 5.4 – Accuracy progression using the best features of the ranking. The mean accuracy value was determined for both groups, for the moving average procedures with one, two and three trials. The labels with “1trial” correspond to the single trial accuracy classification, and the labels with “2trials” and “3 trials” correspond to the accuracy of the classification with data from the moving average procedures with two and three trials, respectively.

Accuracies of TD individuals are always higher than accuracies of ASD group. However, for each line in the graph, the accuracy values do not change significantly along all the classifications done. Moving average of three trials achieved the best results in both groups, and the opposite happened for single trial classification.

Using, as training data, the features of the first 20 ranking positions, we compared subjects' accuracies for both groups using the non-parametric Wilcoxon rank sum test. We chose the number of best features to use in this analysis by testing different sets and selecting the one that presented relatively higher mean accuracy for both groups in all moving average procedures. Statistical tests showed that accuracies of both groups were statistically different using data from the moving average of one ($p<0.01$), two ($p<0.05$) and three trials ($p<0.02$).

Table 5.2 shows the number of subjects of each group whose accuracy classification (using the 20 best features) was statistically significant ($p<0.05$), obtained from 10000 permutations. Mathematically, let $\widehat{D} = \{D'_1 \dots D'_k\}$ be the set of k randomized versions of the original data D sampled from a given null distribution. The empirical p-value for the classifier f is given as:

$$p = \frac{\left|\{D' \in \widehat{D} : a(f, D') \geq a(f, D)\}\right| + 1}{k + 1} \tag{5.2}$$

in which $a(f, X)$ is the accuracy value of the application of the classifier f in a given dataset X (Ojala and Garriga, 2009).

It is clear that TD group had always more subjects with statistically significant classification accuracies than ASD group. The lowest accuracy value belongs to ASD individuals (40 %), and the highest, 73 %, belongs to TD individuals.

Table 5.2 – Number of subjects with statistically significant accuracy classification for ASD and TD groups, and for moving average procedures of one, two and three trials.

Trials		Number of Subjects (p<0.05)		Number of Subjects (p<0.05)
1	**TD**	8/17	**ASD**	5/17
2		8/17		3/17
3		11/17		3/17

5.1.1.1 Mental Imagery of Happy Facial Expression *versus* No-Imagery

In this analysis, we tried to differentiate between mental imagery states from no-imagery states, but using only the trials corresponding to the mental imagery of happy facial expression. The best classification results were obtained using the other baseline interval tested, from 450 ms to 0 ms pre-stimulus.

Similarly to what was described earlier, after obtaining the average ranking positions to all 34 subjects, we determined the frequency bands distribution over the ranking (Figure 5.5). From Figure 5.5 it is possible to see that that theta and alpha frequency bands ([4 8] Hz and [8 15] Hz, respectively) have the lowest mean ranking positions; bands with highest frequency have also higher ranking positions and, so, they are expected to contribute less to better classification results.

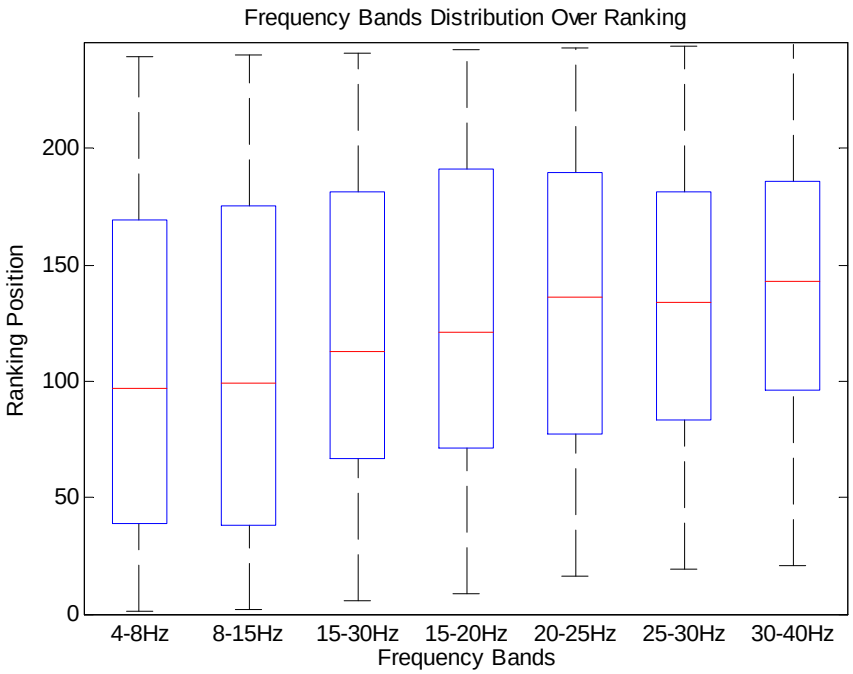

Figure 5.5 – Distribution of the different frequency bands over ranking. Ranking positions of the features corresponding to each frequency band were plotted in this box plot.

From Figure 5.6 and Figure 5.7, which correspond to time-domain and non-linear features distribution, respectively, it is possible to see that non-linear features, in opposite to time-domain features, have their mean ranking positions above the 150th position, for all frequency bands. This fact supports the hypothesis that non-linear features do not have much relevance in the classification of the two conditions analyzed. The outlier feature that appears in Figure 5.7 in each frequency band corresponds to the Correlation Dimension.

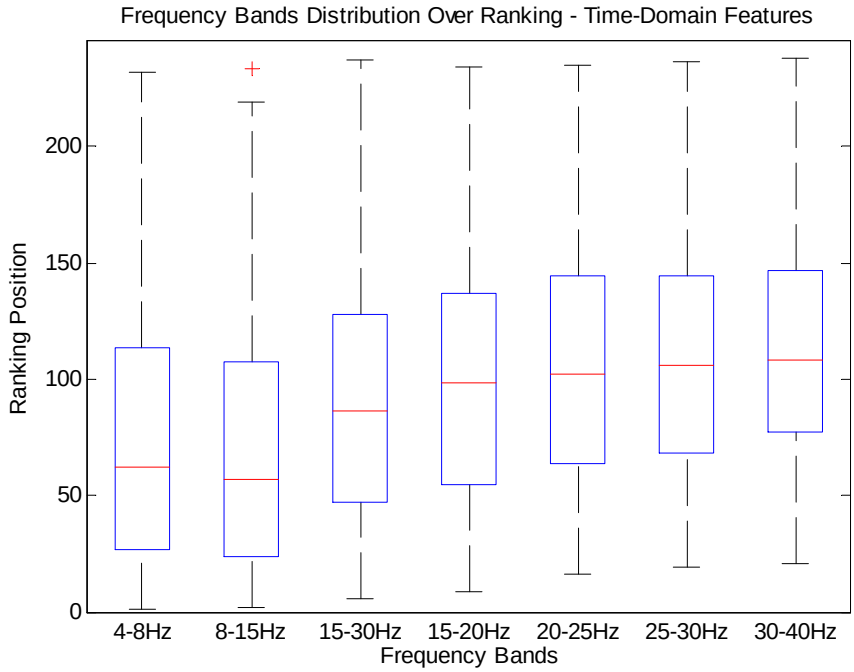

Figure 5.6 – Distribution of the different frequency bands over ranking. Ranking positions of the time-domain features corresponding to each frequency band were saved and plotted in a box plot graph.

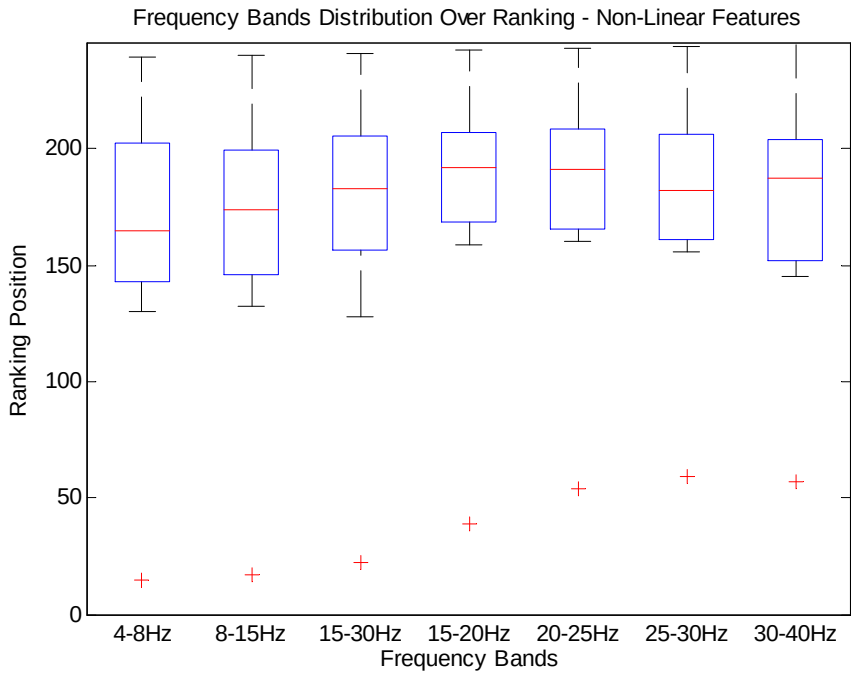

Figure 5.7 – Distribution of the different frequency bands over ranking. Ranking positions of the time-domain features corresponding to each frequency band were saved and plotted in a box plot graph.

Table 5.3 presents the best 10 features of the ranking, determined for the classification between imagery of happy facial expressions and no-imagery epochs. Absolute time-domain features from theta and alpha bands appear to be the most relevant among this set of features; no non-linear features were present in this set of the best 10 features.

Table 5.3 – Characteristics of the 10 best features of the ranking for this analysis.

Ranking Position	Feature				Frequency Band (Hz)
	Time-Domain		Non-Linear		
	Absolute	Baseline Correction	Absolute	Baseline Correction	
1	Average Env	-	-	-	[4 8]
2	Average Env	-	-	-	[8 15]
3	SD Env	-	-	-	[4 8]
4	SD Env	-	-	-	[8 15]
5	Maximum Env	-	-	-	[4 8]
6	Average Env	-	-	-	[15 30]
7	Average Pow	-	-	-	[8 15]
8	Maximum Env	-	-	-	[8 15]
9	Average Env	-	-	-	[15 20]
10	Average Pow	-	-	-	[4 8]

We built the linear SVM model for classification procedure using the corresponding data from the best features of the average ranking. Classification accuracies were determined, also, through a 5-fold cross validation technique. We performed that procedure was performed over the data resultant from the moving average of one, two and three trials, as explained. Figure 5.8 shows the accuracy progression using from 1 to the 8 best features.

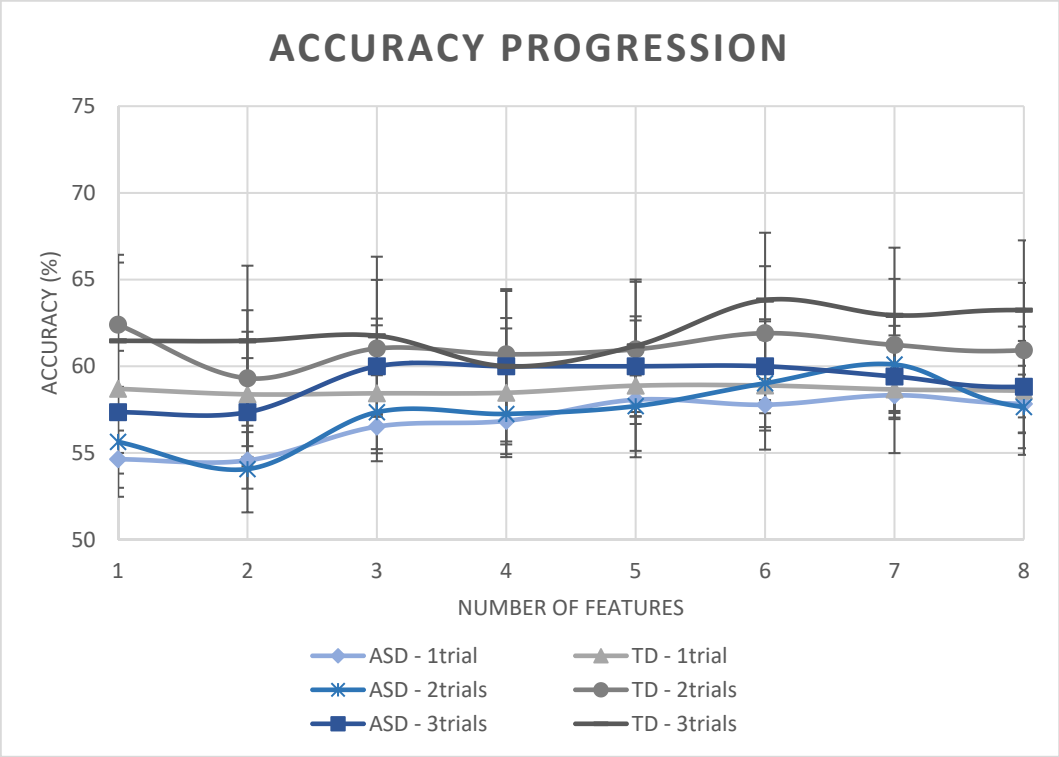

Figure 5.8 – Accuracy progression using the best features of the ranking. The mean accuracy value was determined for both groups, for the moving average procedures with one, two and three trials. The labels with “1trial” corresponds to the single trial accuracy classification, and the labels with “2trials” and “3 trials” correspond to the accuracy of the classification with data from the moving average procedures with two and three trials, respectively. No convergence was achieved within the maximum number of iterations from more than 8 features. Accuracy progression was determined using the mean accuracy of the group and the correspondent SEM for each feature set.

Classification results were similar for both groups; however, accuracies of TD groups were slightly higher. The range of accuracies within each group was high leading to high values of SEM, particularly in TD group. Classification with the moving average of three trials originated the highest accuracy values in each group and, also, the highest SEM values.

We performed a non-parametric Wilcoxon rank sum test to determine if there were statistical differences between the accuracies of both groups. This statistical test was performed using the classification results of each subject using the six best features. We also chose the number of features to use as the one that appear to have relatively high mean accuracy in both groups and moving average approaches. Statistical tests showed that accuracies of both groups were not statistically different for none of the moving average datasets.

Table 5.4 – Number of subjects with statistically significant accuracy classification for ASD and TD groups, and for moving average procedures of one, two and three trials.

Trials		Number of Subjects (p<0.05)		Number of Subjects (p<0.05)
1	**TD**	6/17	**ASD**	10/17
2		8/17		7/17
3		9/17		7/17

Table 5.4 shows the number of subjects of each group whose accuracy classification were statistically significant ($p<0.05$) obtained from the 10000 permutations (Equation (5.2)), using the same six best features. Except for single trial classification, the number of subjects with statistically significant classification accuracy was slightly higher in TD group than in ASD group. The highest accuracy value was 100 % obtained for a TD individual in all moving average trials' accuracy. TD individuals' range of accuracies was from 40 % to 100 % and the range of accuracies of ASD individuals was from 40 % to 85 %.

5.1.1.2 Mental Imagery of Sad Facial Expression *versus* No-Imagery

As performed for mental imagery of happy facial expression, in this analysis we tried to differentiate between mental imagery states from no-imagery states, but using only the trials corresponding to the mental imagery of sad facial expression. Like for mental imagery of happy facial expressions, we obtained the best classification results using the time points from 450 ms to 0 ms pre-stimulus as baseline interval.

We also determined frequency band distribution over ranking positions, using all features extracted (Figure 5.9), only time-domain features (Figure 5.10) and only non-linear features (Figure 5.11).

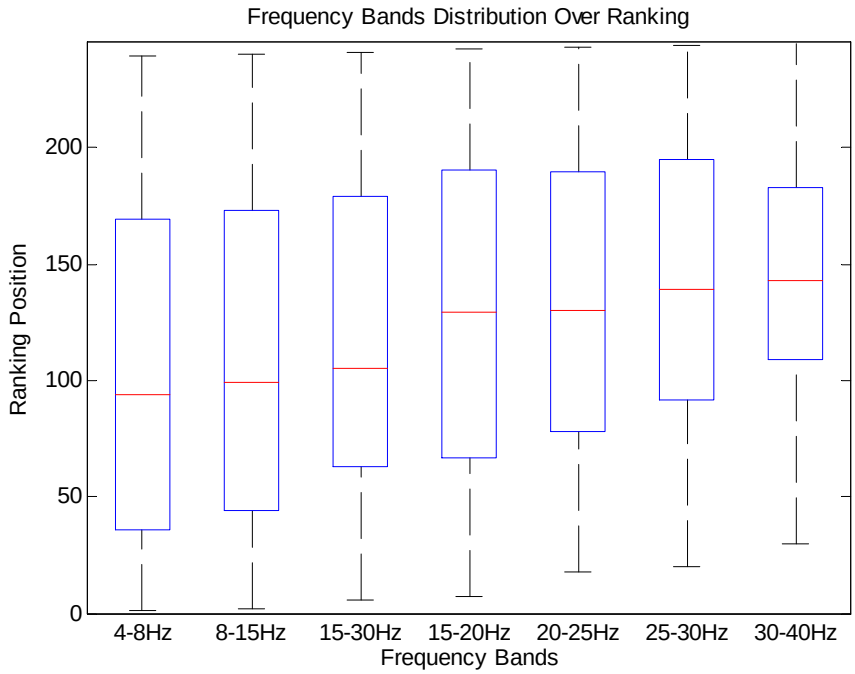

Figure 5.9 – Distribution of the different frequency bands over ranking. Ranking positions of the features corresponding to each frequency band were plotted in this box plot.

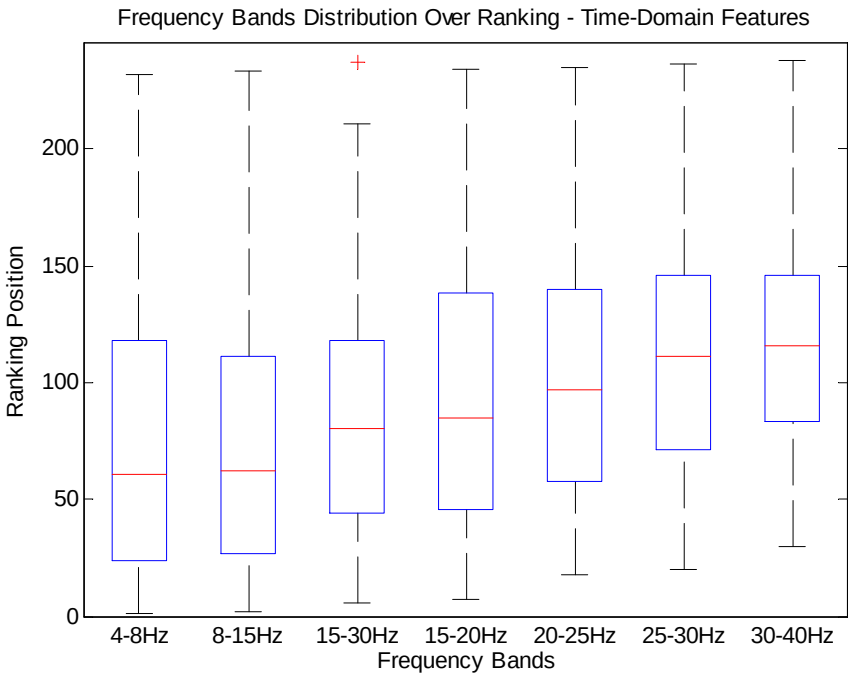

Figure 5.10 – Distribution of the different frequency bands over ranking. Ranking positions of the time-domain features corresponding to each frequency band were saved and plotted in a box plot graph.

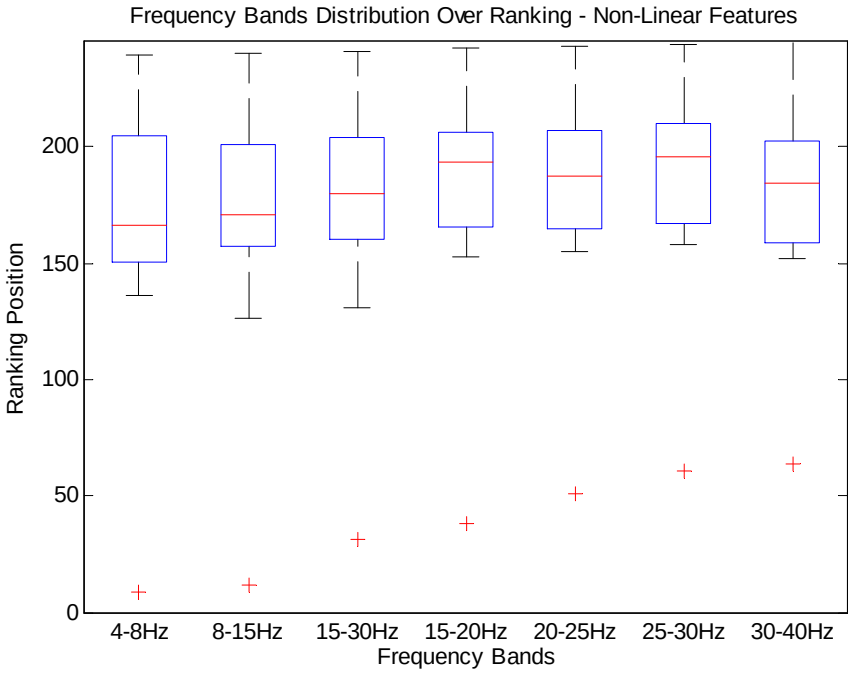

Figure 5.11 – Distribution of the different frequency bands over ranking. Ranking positions of the non-linear features corresponding to each frequency band were saved and plotted in a box plot graph.

Results were very similar to the results obtained for happy mental imagery *versus* no imagery. Theta and alpha frequency bands appear to be more frequently at the lowest positions of the ranking than high frequency bands, exhibiting a large contribution to the classification accuracies with low number of features. As previously reported, non-linear features appear always in the last positions of the ranking, do not showing great contribution to the achievement of high accuracy values. Like in Figure 5.7, the outlier feature that appears in Figure 5.11 in each frequency band corresponds to the Correlation Dimension.

As done in the previous analysis, the 10 best features of the ranking are presented in Table 5.5. Also, as Figure 5.9, Figure 5.10 and Figure 5.11 show, alpha and theta bands were the most common frequency bands in the set of the 10 features with the lowest ranking values. The dominance of the time-domain features in the lowest positions of the ranking is also notorious. It should be noted also that except the feature Maximum Env of [8 15] Hz band, at happy facial expression imagery *versus* no imagery, and feature CorrDim of [4 8] Hz band, in this analysis all the other features of the 10 best features set are the same in both analysis, with slightly different ranking positions.

Table 5.5 – Characteristics of the 10 best features of the ranking for mental imagery of sad facial expression versus no-imagery analysis.

Ranking Position	Feature				Frequency Band (Hz)
	Time-Domain		Non-Linear		
	Absolute	Baseline Correction	Absolute	Baseline Correction	
1	Average Env	-	-	-	[4 8]
2	Average Env	-	-	-	[8 15]
3	SD Env	-	-	-	[4 8]
4	SD Env	-	-	-	[8 15]
5	Maximum Env	-	-	-	[4 8]
6	Average Pow	-	-	-	[4 8]
7	Average Env	-	-	-	[15 30]
8	Average Env	-	-	-	[15 20]
9	Average Pow	-	-	-	[8 15]
10	-	-	CorrDim	-	[4 8]

Also, we used a linear SVM model for the classification procedure. In Figure 5.12 it is possible to see a clear gap between the accuracies of TD and ASD groups; classification accuracies of TD individuals achieved much higher values than the ones achieved by ASD group. In TD group, single trial classification obtained the lowest accuracy values and classification with moving average of three trials achieved the highest. In ASD group, these differences were not clear. As reported in the previous analysis, with the happy facial expression, the range of accuracies within each group is very high, especially in TD group, leading to higher SEM.

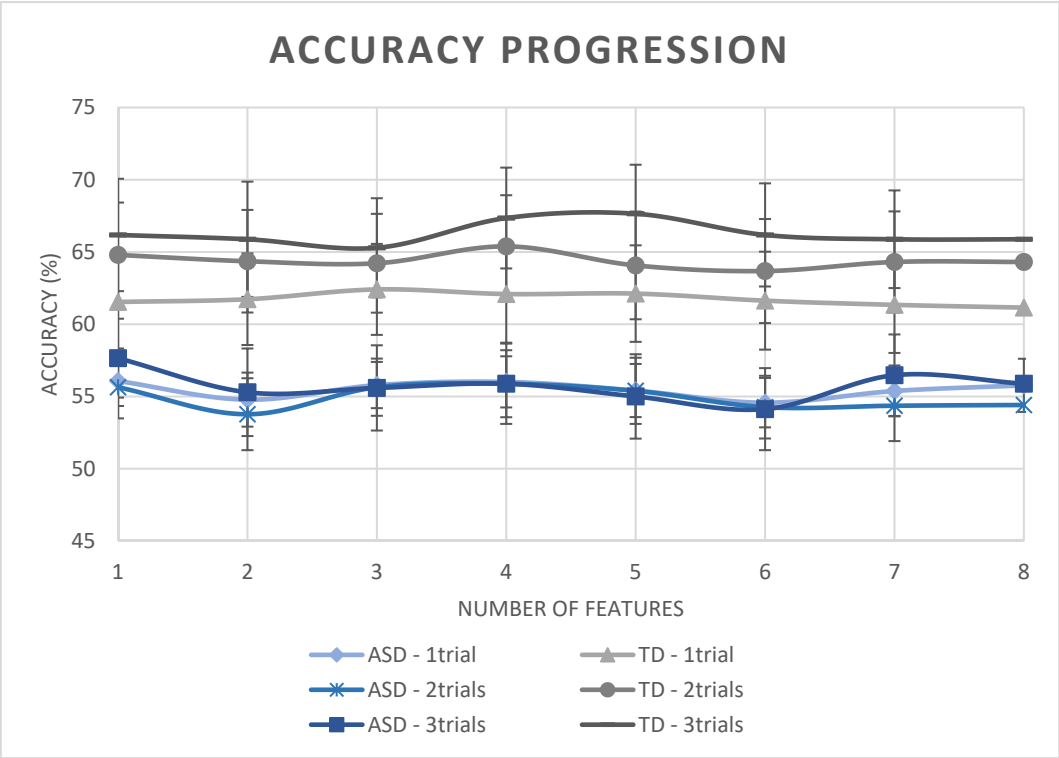

Figure 5.12 – Accuracy progression using the best features of the ranking. The mean accuracy value was determined for both groups, for the moving average procedures with one, two and three trials. The labels with “1trial” corresponds to the single trial accuracy classification, and the labels with “2trials” and “3 trials” correspond to the accuracy of the classification with data from the moving average procedures with two and three trials, respectively. The linear SVM model does not converged with more than 8 features. Accuracy progression was determined using the mean accuracy of the group and the correspondent SEM for each feature set.

We performed a non-parametric Wilcoxon rank sum test in the accuracies results of the subjects of each group using the four best features. We chose this number of best features to use as explained in previous sections. Statistical tests showed statistically significant results for accuracies of the moving average of two and three trials ($p<0.05$). Table 5.6 shows the number of subjects of each group whose accuracy classification were statistically significant ($p<0.05$) obtained from the 10000 permutations (Equation (5.2), using the four best features. The number of subjects with statistically significant classification accuracy was much higher in TD group for the moving average of two and three trials. The highest accuracy value was 100 %, obtained for a TD individual in all moving average trials’

accuracy. TD individuals' range of accuracies was from 50 % to 100 % and the range of accuracies of ASD individuals was from 43 % to 80 %.

Table 5.6 – Number of subjects with statistically significant accuracy classification for ASD and TD groups, and for moving average procedures of one, two and three trials.

Trials		Number of Subjects (p<0.05)		Number of Subjects (p<0.05)
1	**TD**	8/17	**ASD**	7/17
2		11/17		4/17
3		12/17		5/17

5.1.2 Mental Imagery of Happy *versus* Sad Facial Expressions

Besides trying to discriminate between mental imagery and no-imagery states, we also tried to discriminate between mental imagery of happy and sad facial expressions, using only the mental imagery time interval. We achieved the best classification results for this analysis using, as baseline, the time interval from 3500 ms to 3000 ms pre-stimulus.

Furthermore, we obtained frequency bands distribution over ranking (Figure 5.13, Figure 5.14 and Figure 5.15). Low frequency bands had also the lowest positions in the ranking, contributing to better accuracy values in the classification process with the best features. The difference in ranking positions of time-domain and non-linear features, although clear, was not so "linear" as in the previous analysis. In time-domain features, it was clear the low median ranking positions for theta and alpha bands but, although low, the median positions of the features from other frequency bands were approximately at the middle of the ranking (Figure 5.14). For non-linear features, in this case, not every frequency band had high mean ranking position, as gamma frequency band presented a median ranking position below the middle of the ranking (Figure 5.15).

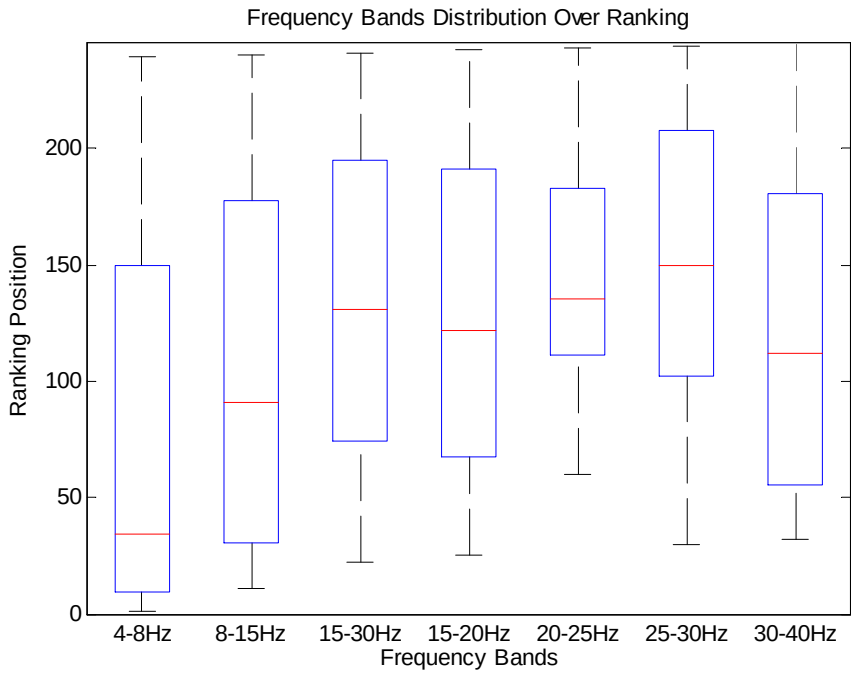

Figure 5.13 – Distribution of the different frequency bands over ranking. Ranking positions of the features corresponding to each frequency band were plotted in this box plot.

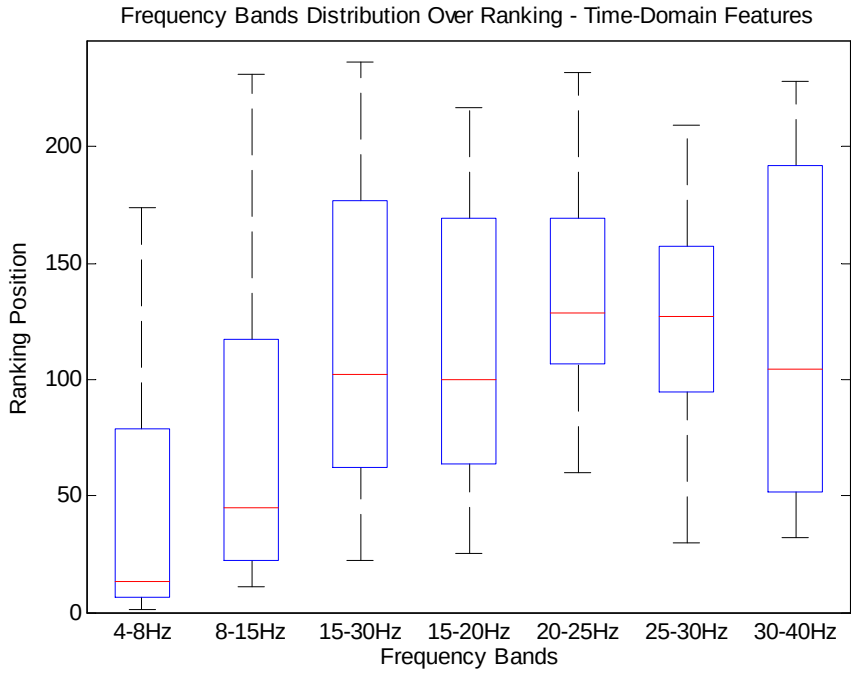

Figure 5.14 – Distribution of the different frequency bands over ranking. Ranking positions of the time-domain features corresponding to each frequency band were saved and plotted in a box plot graph.

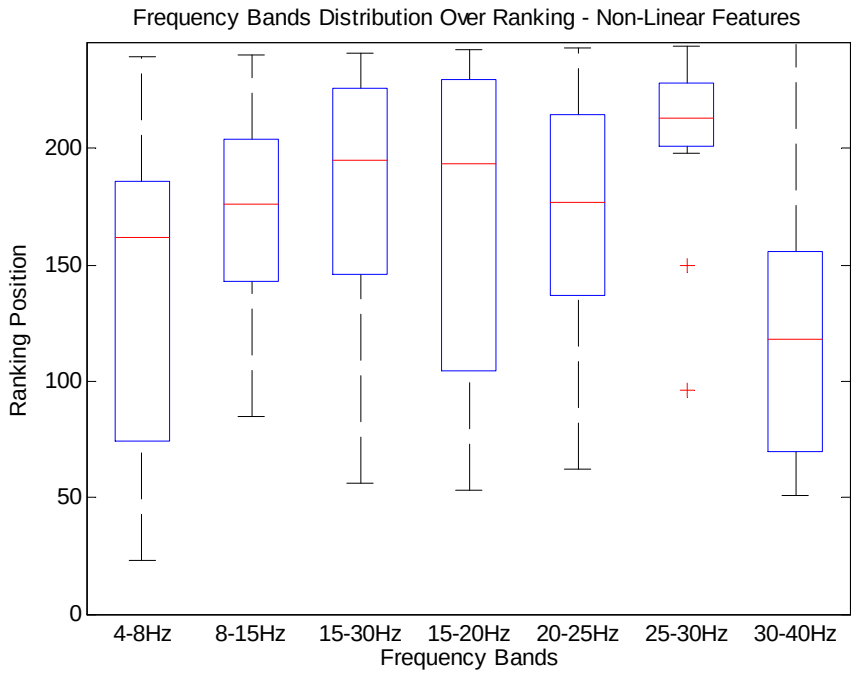

Figure 5.15 – Distribution of the different frequency bands over ranking. Ranking positions of the non-linear features corresponding to each frequency band were saved and plotted in a box plot graph.

In Table 5.7, which contains the characteristics of the 10 best features, it is possible to see that all the features were extracted from the theta frequency band, which had the lowest mean ranking position. The features were all, also, from time-domain.

Table 5.7 – Characteristics of the 10 best features of the ranking for discrimination of mental imagery of happy and sad facial expressions.

Ranking Position	Feature				Frequency Band (Hz)
	Time-Domain		Non-Linear		
	Absolute	Baseline Correction	Absolute	Baseline Correction	
1	Average Teag	-	-	-	[4 8]
2	Maximum Teag	-	-	-	[4 8]
3	SD Pow	-	-	-	[4 8]
4	-	Maximum Power	-	-	[4 8]
5	SD Env	-	-	-	[4 8]
6	Maximum Pow	-	-	-	[4 8]
7	SD Teag	-	-	-	[4 8]
8	Average Pow	-	-	-	[4 8]
9	Maximum Env	-	-	-	[4 8]
10	Average Env	-	-	-	[4 8]

Figure 5.16 presents the progression of the mean accuracy values, for each group, using data from 5 to the 50 best features of the ranking. We determined accuracy values resultant from the linear SVM classifier using a 5-fold cross validation technique.

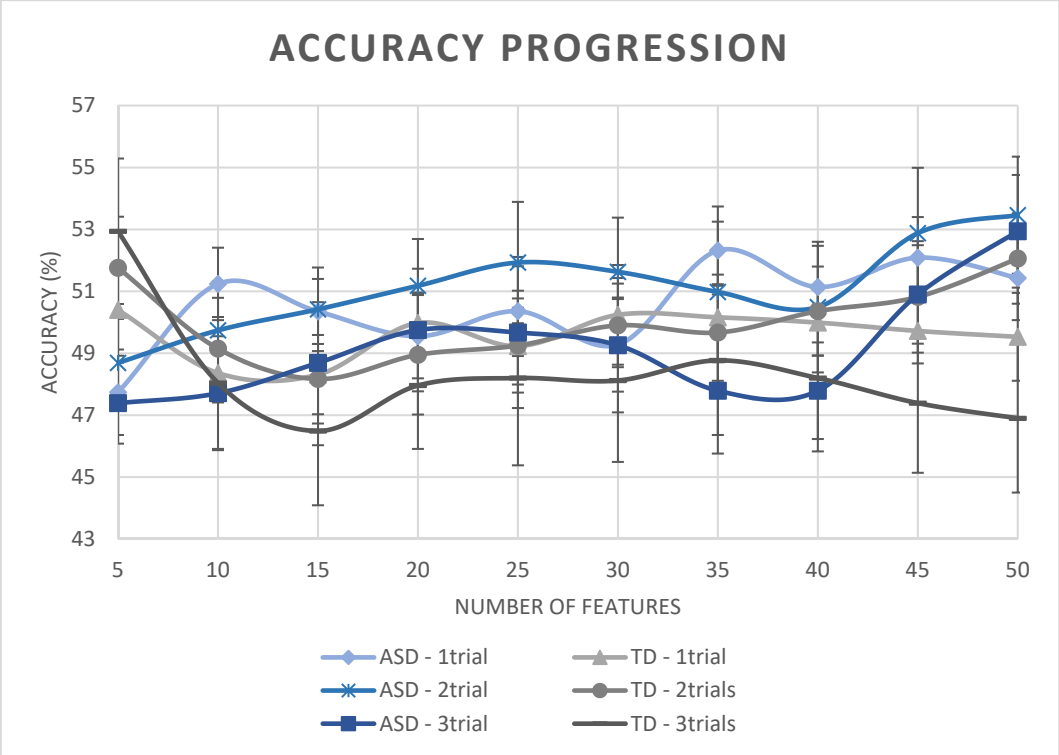

Figure 5.16 – Accuracy progression using the best features of the ranking. The mean accuracy value was determined for both groups, for the moving average of one, two and three trials. The labels with “1trial” corresponds to the single trial accuracy classification, and the labels with “2trials” and “3 trials” correspond to the accuracy of the classification with data from the moving average procedures with 2 and 3 trials, respectively.

In this case, accuracy values were slightly higher in ASD group, and did not differ very much regarding the moving average procedure.

We performed statistical non-parametric Wilcoxon rank sum test in accuracy values of the subjects from both groups, using only the five best features. Mean group accuracy appeared to change randomly with the number of features used for classification; thus, we chose the five first features to apply the statistical tests. Results revealed a statistical significant difference only in the moving average of three trials ($p<0.05$). Using the same dataset, we also determined the number of subjects with statistically significant classification accuracy. We achieved the statistical p-value from 10000 permutations, according to Equation (5.2). The

results are shown in Table 5.8. No subjects from TD group showed statistically significant accuracy classification and few from ASD group did.

Table 5.8 – Number of subjects with statistically significant accuracy classification for ASD and TD groups, and for the moving average of one, two and three trials.

Trials		Number of Subjects (p<0.05)		Number of Subjects (p<0.05)
1	**TD**	0/17	**ASD**	1/17
2		0/17		1/17
3		0/17		4/17

5.2 Classification of Autism Spectrum Disorder versus Typically Developing Individuals

This analysis is completely different from the ones explained in previous sections. Here, we tried to find biomarkers that could be used to distinguish ASD individuals from TD individuals. To do that, we extracted 35 features from each EEG channel signal filtered for seven different frequency bands, during the imagery epoch [500 4000] ms. We extracted the baseline epoch from 3500 ms to 3000 ms pre stimulus.

Using the information given by the ranking of the features extracted, it was possible to analyze the distribution of the features from each frequency band through the ranking. We performed that analysis through visual inspection of the box plot distribution (Figure 5.17). Gamma, High Beta ([25 30] Hz) and theta frequency bands showed the lowest median ranking positions. We expected that most of the features from those frequency bands to be present at the lowest

positions of the ranking and highly relevance for the discrimination between the two groups of individuals.

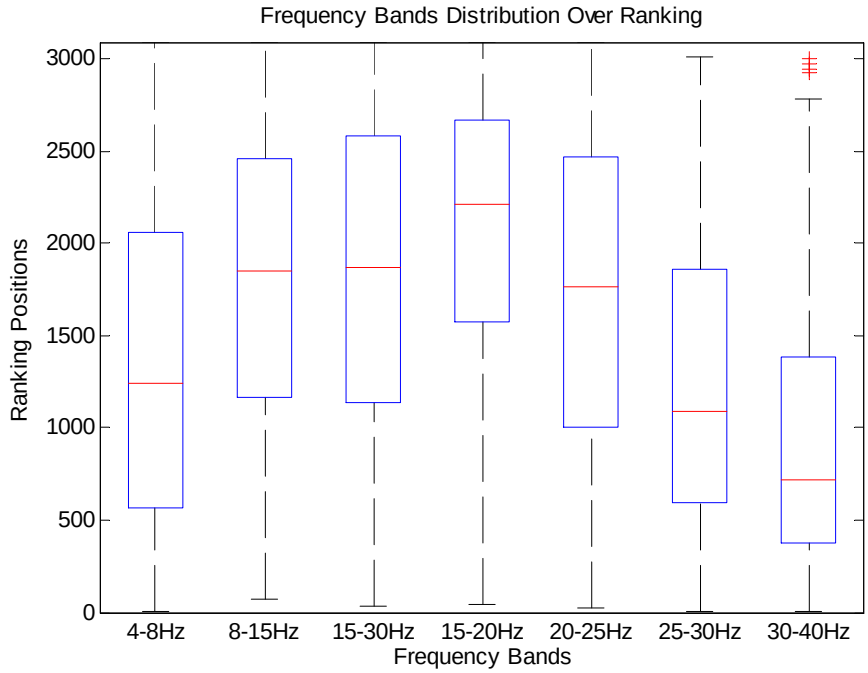

Figure 5.17 – Distribution of the different frequency bands over ranking. Ranking positions of each feature, from each channel cluster, corresponding to the frequency bands studied was determined. The distribution of the ranking positions of the features belonging to each frequency band were plotted in this box plot.

We also determined frequency distribution over ranking for time-domain and non-linear features (Figure 5.18 and Figure 5.19, respectively).

Time-domain features appear to occupy higher ranking positions than non-linear features for almost frequency bands studied. It should be noted that theta, high beta and gamma frequency bands occupy also the lowest mean ranking positions. Nevertheless, theta band appears to be more relevant in non-linear features, while gamma and high beta frequency bands appear to have more importance in time-domain features.

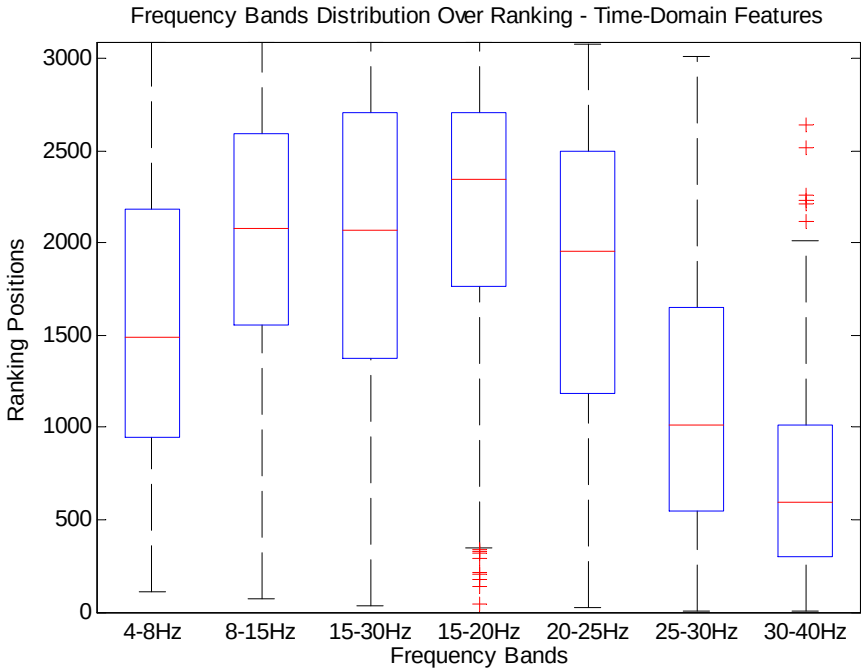

Figure 5.18 – Distribution of the different frequency bands over ranking. Ranking positions of each time-domain feature corresponding to each frequency bands studied was determined. The distribution of the ranking positions of those features, belonging to each frequency band, were plotted in this box plot.

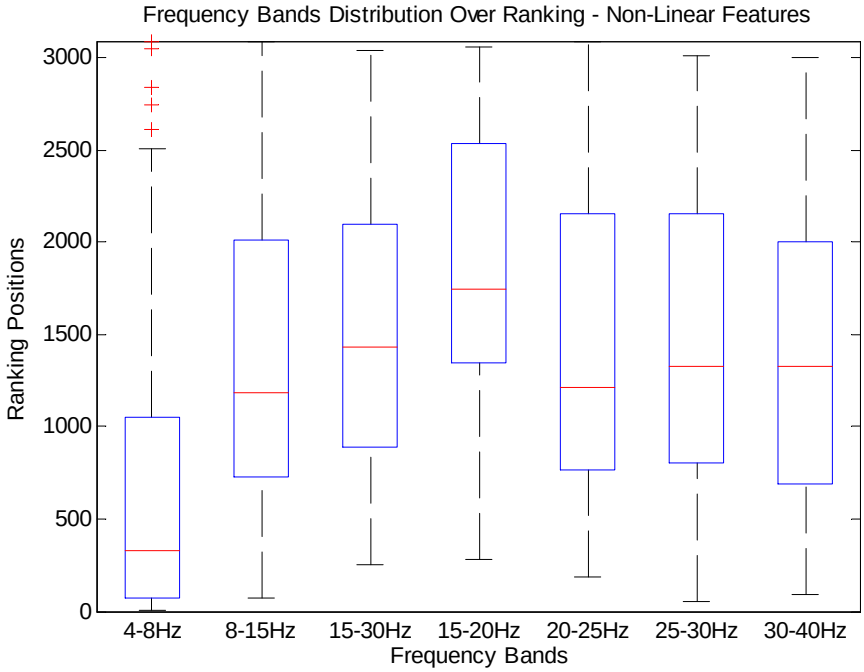

Figure 5.19 – Distribution of the different frequency bands over ranking. Ranking positions of each non-linear feature corresponding to each frequency bands studied was determined. The distribution of the ranking positions of those features, belonging to each frequency band, were plotted in this box plot.

We also analyzed the ranking distribution over the channel clusters (Figure 5.20).

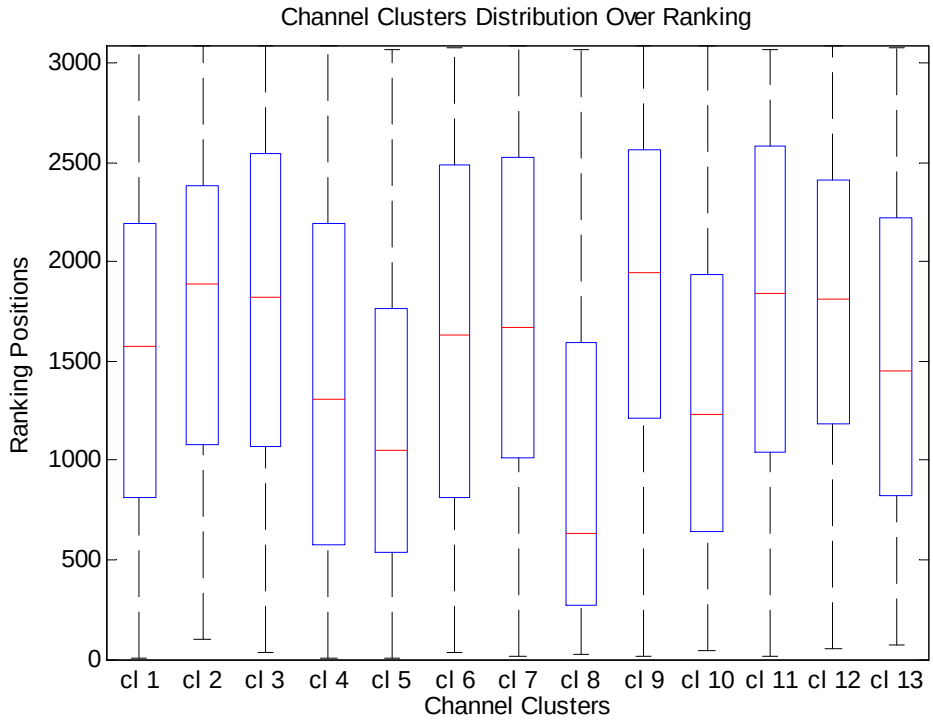

Figure 5.20 – Distribution of the ranking positions of the features belonging to the different channel clusters defined. Ranking positions of each feature corresponding to each channel cluster was determined. The distribution of the ranking positions of those features, belonging to each channel cluster, were plotted in this box plot.

Channel clusters that appear to have more influence for distinguishing between the two groups are Cluster 8, Cluster 5, Cluster 10 and Cluster 4, corresponding to, respectively, right Fronto-Temporal, right Centro-Parietal, left Centro-Parietal, and right Parieto-Occipital regions.

Table 5.9 presents the 10 best features according to FS ranking. Those features were mostly from time-domain and baseline corrected values of the features extracted. All time-domain features of this set were estimated from [25 30] Hz and, in smaller number, from [30 40] Hz frequency bands, as expected from the visual analysis of Figure 5.18. The only three non-linear features present in this set were extracted from the lowest frequency band, theta band ([4 8] Hz), also visible

through the analysis of Figure 5.19. The most frequent cluster was the fourth, corresponding to electrodes at the Occipital and Parieto-Occipital sites in the right hemisphere; this cluster had one of the lowest mean positions in the ranking, showed in Figure 5.20.

Table 5.9 – Characteristics of the 10 best features of the ranking for discrimination between ASD and TD individuals.

Ranking Position	Feature				Frequency Band (Hz)	Channel Cluster
	Time-Domain		Non-Linear			
	Absolute	Baseline Correction	Absolute	Baseline Correction		
1	-	Average Env	-	-	[25 30]	4
2	-	Average Env	-	-	[30 40]	4
3	-	-	-	SFI	[4 8]	4
4	-	Average Teag	-	-	[25 30]	1
5	-	Average Pow	-	-	[25 30]	4
6	-	Average Env	-	-	[30 40]	5
7	-	Average Teag	-	-	[25 30]	4
8	-	-	-	Lyap	[4 8]	4
9	-	Average Pow	-	-	[25 30]	1
10	-	-	-	SFI	[4 8]	11

After knowing which were the best features, we classified the data using a certain number of the best features, ordered by their ranking position. Starting with the five best features from the ranking, we repeated the classification process by adding the next five features until reaching a set with the 100 best features. Figure 5.21 presents the accuracy progression of the classification accuracy values obtained using from 5 to 100 best features. These values correspond to the leave-one-out accuracy classification of each subject.

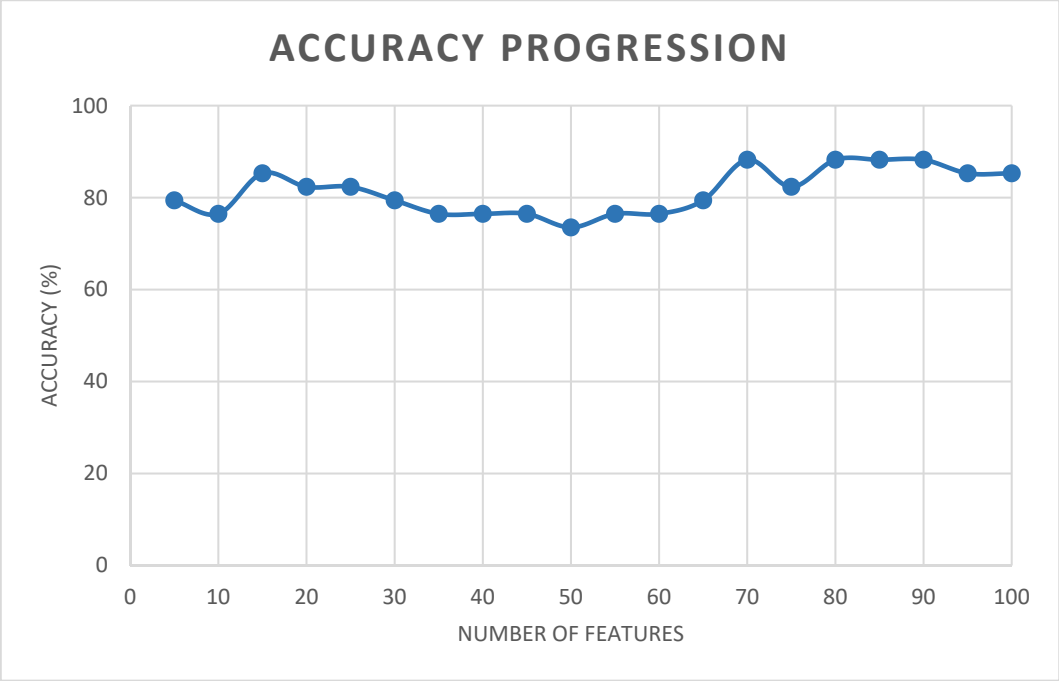

Figure 5.21 – Progression of the accuracy classification values, obtained using from 5 to the 100 best features of the ranking.

In order to reduce even more the data dimensionality, after selecting the data corresponding to the 100 best features, we applied a PCA analysis to this data using two different approaches (explained in section 4.5.2.2).

In the first approach, we directly applied the PCA analysis to the data from 100 best features given from the FS algorithm to reduce redundancy between features.

Then, using the leave-one-out cross validation method, we classified the data using until its 15 principal components, and we built a graph with the accuracy progression (Figure 5.22).

Classification accuracy seems to stabilize at a very high accuracy value from the first five principal components.

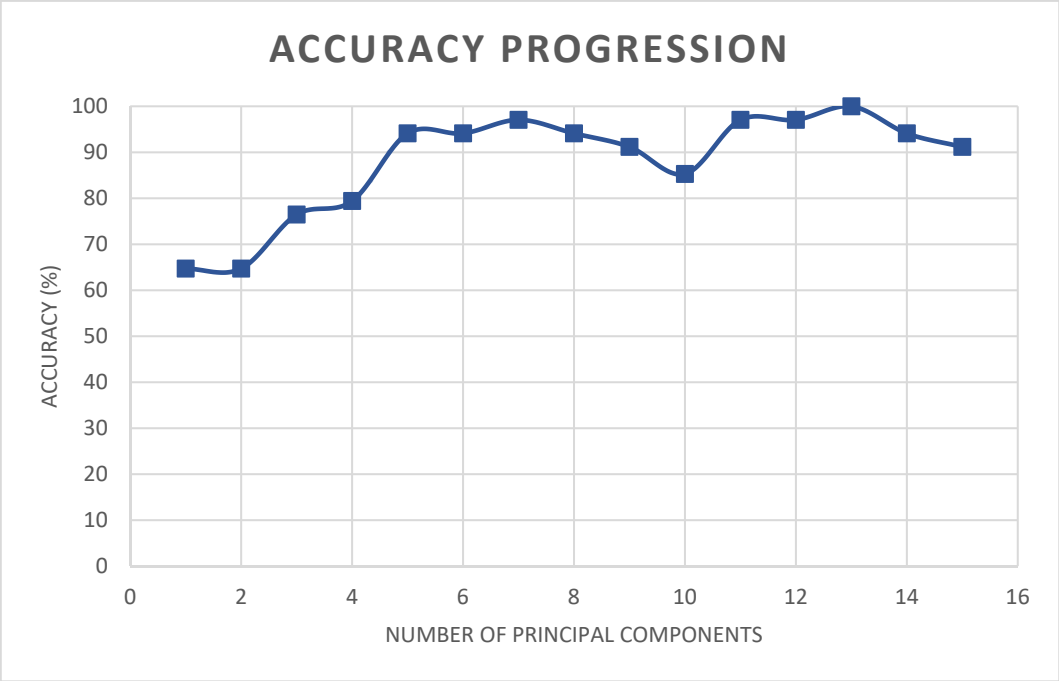

Figure 5.22 – Progression of the accuracy classification values of the first approach of the analysis. Accuracy values were determined using up to the 15 first principal components, resultant from the application of a PCA analysis to the data corresponding to the 100 best features of the ranking.

In the second approach, through visual inspection of the Figure 5.17, Figure 5.18, Figure 5.19, and Figure 5.20, we selected the channel clusters and frequency bands which had the lowest median ranking positions. We selected four channel clusters, Cluster 8, Cluster 5, Cluster 10 and Cluster 4 (represented in red circles – Figure 5.23), as well as three frequency bands, theta ([4 8] Hz), sub-band high beta ([25 30] Hz), and gamma ([30 40] Hz).

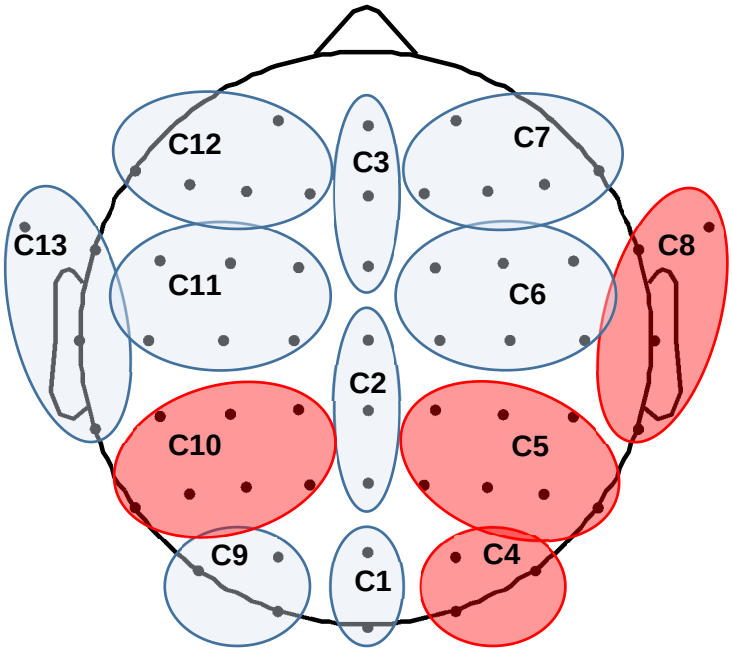

Figure 5.23 – Representation of the locations of the channel clusters selected. Clusters 8, 5, 10 and 4 correspond, respectively, to right Fronto-Temporal, right Centro-Parietal, left Centro-Parietal, and right Parieto-Occipital electrode positions.

Figure 5.24 shows the distribution of the frequency bands through the ranking relative to the features of each one of the four clusters selected. From the box plot distributions, it is clear that the frequency bands selected had the lowest median ranking positions for the features of the four channel clusters selected.

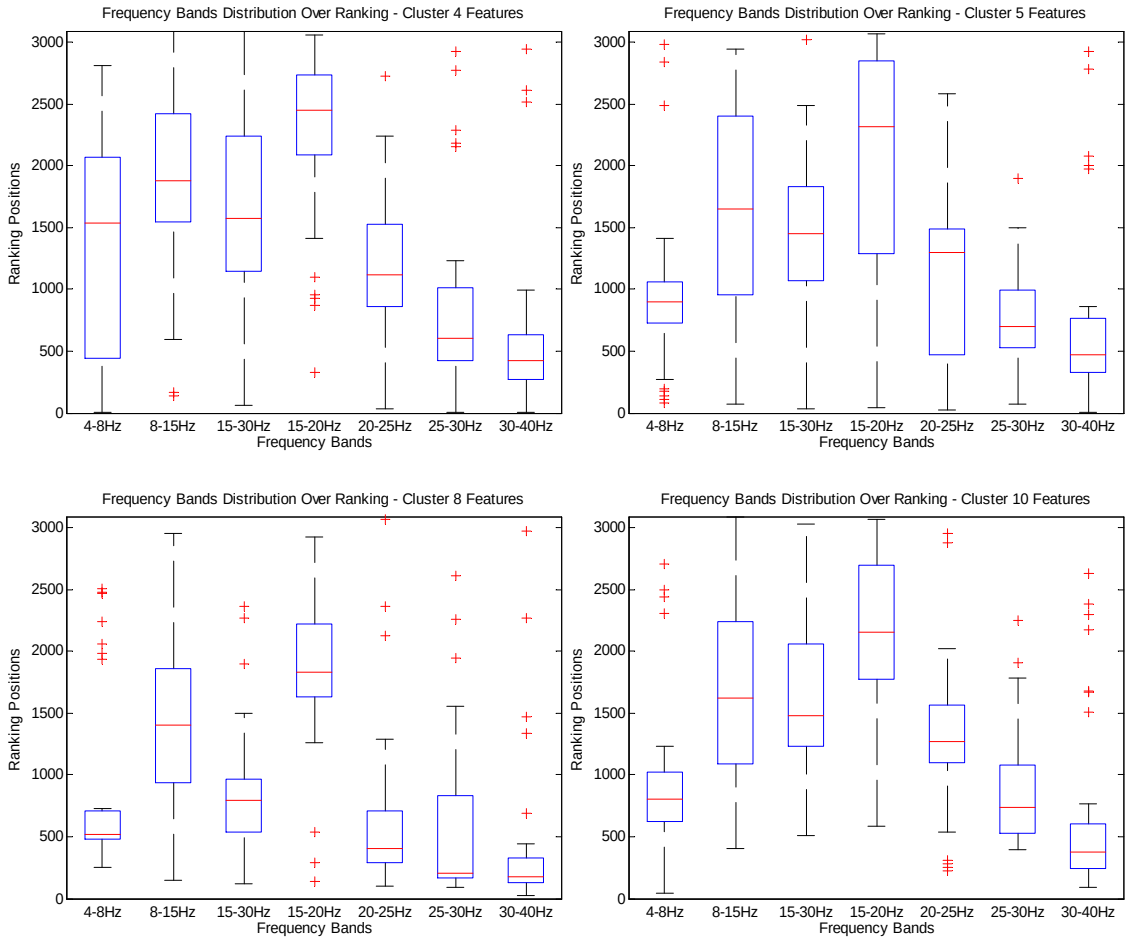

Figure 5.24 – Distribution of the ranking positions of the features corresponding to each one of the four clusters selected through each frequency band. In all four clusters, it is visible that frequency bands with the lowest mean ranking positions are theta, high-beta and gamma. The cluster 8 appears to have the lowest mean ranking positions in all frequency bands, which corroborates the ranking position distribution in Figure 5.20.

Then, to reduce even more the dimensionality and to narrow the range of features to use in classification process, we extracted, from the 100 best features data, the features corresponding to these selected channel clusters and frequency bands. The final data, after the selection, had 31 features, being that 13 were related to [4 8] Hz frequency band, 7 to [25 30] Hz and 11 to [30 40] Hz; all the features from theta band were non-linear and those from the other two frequency bands selected were mostly time-domain features (as expected from the analysis of Figure 5.18 and Figure 5.19); cluster 4 had the most contributions of features in

this set, followed by cluster 10, 8 and 5 respectively. To these data, with the 31 selected features, the PCA analysis was done. We used the principal components to classify the data using the leave-one-out cross validation technique. Figure 5.25 shows the progression of accuracy classification values using from 1 to the 15 principal components resultant from PCA analysis.

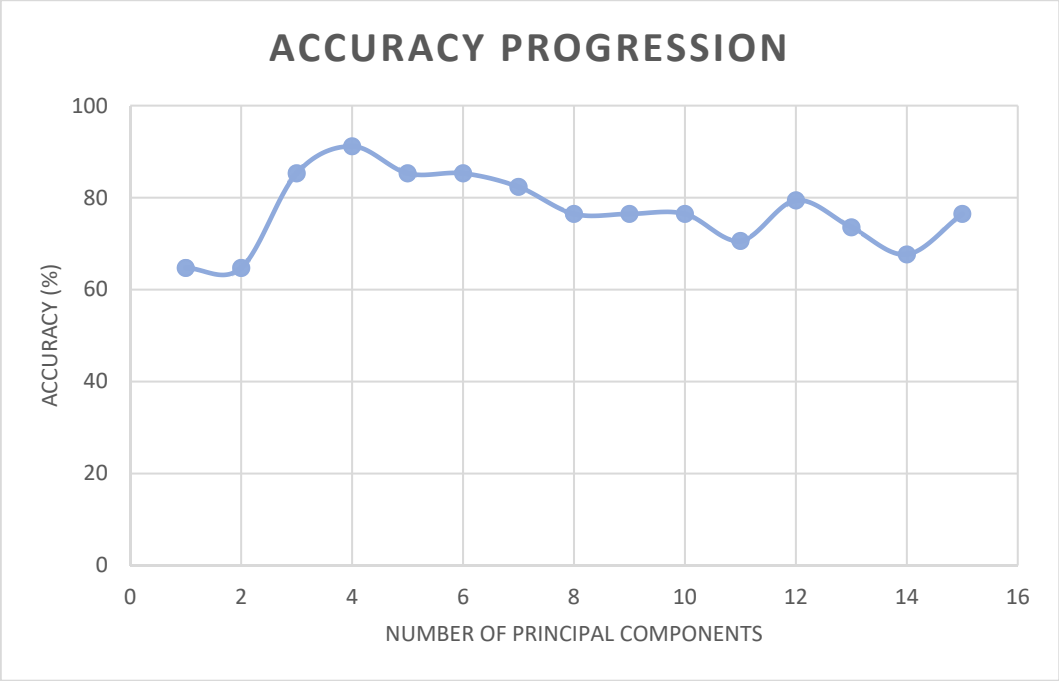

Figure 5.25 – Progression of the accuracy classification values of the second approach of the analysis. Accuracy values were determined using up to the 15 first principal components, resultant from the application of a PCA analysis to the data corresponding to the features of the best clusters and frequency bands (selected from the 100 best features of the ranking).

Comparing the accuracy progressions obtained in both approaches (see Figure 5.26), it is clear that the second approach had high values of accuracy using less than the 5 first principal components. The first approach resulted in the best accuracy values for almost all of the sets of principal components tested. The second approach revealed to be very restrictive in the features to use for PCA, leading to, although high, not so good results.

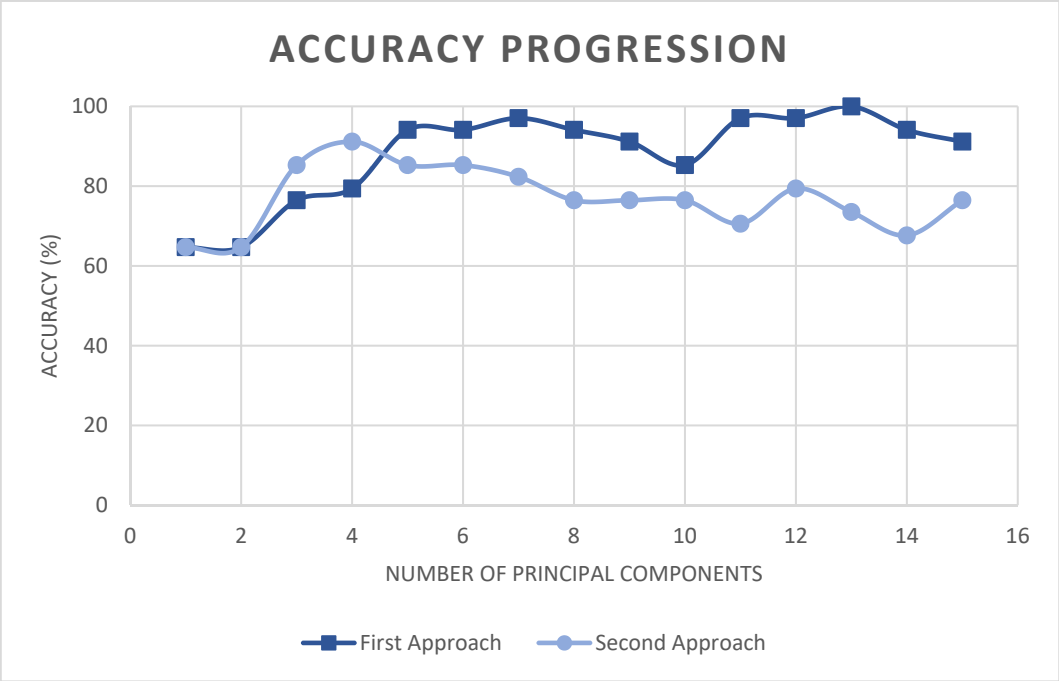

Figure 5.26 – Comparison of the two approaches tested for the discrimination between ASD and TD individuals. The highest accuracy value obtained using the second approach was of 91.17 %, using 4 principal components. Using 5 principal components, accuracy using the first approach was of 94.12 % and, accuracy resultant from the second approach was of 85.29 %, and using 10 principal components the accuracy values obtained were 85.29 % and 76.47 %, respectively.

Chapter 6

Discussion and Conclusions

The main objective of this work was to evaluate the viability of the development of an EEG-based BCI with neurofeedback to rehabilitate social cognitive impairments in ASD individuals through the training of the processing and mental imagery of facial expressions. To do that, we implemented different classifier approaches and tested different conditions: firstly, to distinguish between the mental imagery of the two different facial expressions and, at last, to distinguish between the two groups of individuals and then to distinguish between mental imagery times and no imagery times.

6.1 Classification of Mental Imagery of Facial Expressions

Initially, we performed classification tests using the baseline interval from 450 ms to 0 ms pre-stimulus, which corresponds to a small epoch of the preparation time (after the visual instruction and before the imagery period). As the classification results obtained were not very good, around 50 %, possibly because the imagery networks were already being recruited, we discussed a new baseline interval, before the visual instruction time, from 3500 ms to 3000 ms pre-beep. This is because as individuals received the instruction before the beginning of each run, they could have started a mental imagery preparation after visual instruction. With this new baseline, the classification results were slightly better, even though not above 55 % to happy *versus* sad imagery and 60 % for facial expressions imagery

versus no imagery. For the classification of happy mental imagery *versus* no imagery and sad mental imagery *versus* no imagery the results were nevertheless better with the first baseline interval defined, between visual instruction and mental imagery times.

Regarding the distribution of the features in the ranking, obtained through the application of the Fisher Score algorithm to the training data, in all four analysis performed, it was clear the dominance of time-domain features at the first positions of the ranking; notice that the first ranked features are the ones with the highest relevance to discriminate between the two conditions that are being analyzed. Time-domain features appear to be more relevant than non-linear features in order to differentiate imagery from no imagery times. In all the analysis performed, the median position of the ranking distribution of non-linear features appears always above the middle ranking position. In the first analysis, where it is compared the mental imagery of facial expressions (happy and sad) with no-imagery times, it is possible to see clearly (Figure 5.2) that theta and gamma frequency bands had the lowest median distribution position in the ranking. In every other analysis performed, the dominant frequency bands were clearly theta and alpha, as can also be seen in the analysis of the tables containing the 10 best features. Theta rhythm is related to processing of emotions, focused concentration during mental tasks and imagery states (Collura, 1997; Tatum, 2014), and alpha rhythm is mostly recorded over occipital and parietal regions (Clark, 2009; Nunez and Srinivasan, 2006; Tatum, 2014), which correspond to regions where the core system for face perception is located.

As Fisher Score algorithm choses the features independently, and not the set of features that better separate the two conditions, by classifying with, for example, the best 10 features of the ranking, it does not mean that this specific set would give the better results. Also, that can explain the fact that accuracy values do not stabilize with the increase of the number of features. However, it was necessary to apply an algorithm that minimally ranked the features extracted in order to reduce the dimensionality of the dataset.

In the first three analyses (sections 5.1.1, 5.1.1.1 and 5.1.1.2), in which we compared mental imagery states with no-imagery states, TD group had always slightly better accuracy results than ASD group. Also, we found a higher number

of subjects whose classification was statistically significant ($p<0.05$) for the TD group than for the ASD group.

Comparing the analysis reported in sections 5.1.1.1 and 5.1.1.2, in which the classification of mental imagery of happy or sad facial expressions with no imagery times is performed, it is clear that the first 10 features of the ranking are very similar in both analyses; containing almost the same features although in different ranking positions. An interesting fact is that most of the subjects that had statistically significant classifications in one analysis were the same that had statistically significant accuracy classifications in the other. However, discrimination between imagery of sad facial expressions from no imagery times was easier from TD individuals than the discrimination between imagery of happy facial expressions and no imagery times, and a higher number of individuals had statistical significant classification accuracies. For ASD individuals, mean accuracy group was similar between these two analyses, although the number of individuals with statistically significant classifications was higher in happy mental imagery analysis.

When comparing mental imagery of happy facial expression with mental imagery of sad facial expression (section 5.1.2), the mean group accuracies were slightly higher for ASD than for TD groups, and only a few individuals of ASD group had its classification statistically significant. This last analysis revealed the worst results, as expected, in comparison to the results of the discrimination between “simpler” conditions.

Only for happy versus sad mental imagery discrimination, the moving average over trials did not have any effect. For the other analyses, moving average over three trials resulted in the best accuracy values, for both groups; and the single trial classification resulted in the worst results. Moving average led to the increase of the differences between individuals from the two groups, which can be seen by the statistically significant differences found in the accuracies of subjects from the two groups in the data corresponding to the moving average of two and three trials (in two of the four analyses performed). However, the number of observations to classify in the three trials’ moving average is significantly reduced compared to the single trial classification.

6.2 Classification of ASD versus TD Individuals

In this analysis, the main goal was to find biomarkers that could be used to distinguish ASD from TD individuals.

Frequency bands distribution obtained from Fisher Score feature ranking shows that, in this analysis, not only low frequency bands like theta, but also high frequency rhythms, high-beta and gamma, had the lowest median ranking positions and greatly contribute for the accuracy classification values with low number of features (the best ones). An interesting fact is that theta band presence at the lowest ranking positions is given by the non-linear features, and the presence of the two highest frequency bands studied is given by time-domain features. In this analysis, not only time-domain features are present at the lowest ranking positions, as was the case of imagery vs no-imagery analysis, but also non-linear features were found and at a specific frequency band.

The channel clusters distribution over ranking revealed that Temporal, Parietal and Occipital regions, mostly from the right hemisphere, had great relevance in distinguishing between the two groups. These areas correspond to the location of the core brain regions responsible for face processing, that include, OFA, FFA and, in particular, pSTS, responsible for the processing of variable aspects of the face such as facial expressions.

Accuracy classification values, obtained using from 5 until the 100 best features, was very unstable, and so we decided to do a PCA analysis over the 100 best features data in order to reduce the dimensionality and the redundancy between features. We analyzed the accuracy values using the first 15 principal components and, from the fifth principal component, accuracy values seemed to stabilize around 95 %. In order to improve even more the accuracy values, from the 100 best features we selected only those corresponding to theta, high beta and gamma frequency bands and those from the four best channel clusters. We applied the PCA analysis to those features and then, as previously, we analyzed the accuracy values using from 1 to 15 principal components. However, this selection revealed to be too much restrictive; the accuracy values did not stabilize along the

15 principal components and were higher than in the previous analysis just until the fifth principal component.

6.3 Conclusions

It was possible to distinguish the two groups using the imagery period, however it was not possible to distinguish between states of mental imagery of facial expressions and no-imagery states at an interesting accuracy level.

Results showing that regions from parietal, temporal and occipital right hemisphere greatly contribute to the discrimination between ASD and TD individuals corroborate results previously found by the team (not reported in this dissertation), in which a statistically significant difference between the two groups was found in the right precuneus area for the mental imagery of happy and sad facial expressions.

However, it was not possible to distinguish with relatively high accuracy values the time intervals of mental imagery of facial expressions and no-imagery. The main goal of the project was to evaluate the viability of the development of a EEG-based BCI with neurofeedback approach to rehabilitate ASD individuals, based on this task. With the results obtained it is possible to conclude that the development of that training approach is still impracticable and further study on the subject is needed. Nevertheless, the results also suggest that it is possible do define biomarkers of ASD pathology based on EEG correlates of facial expressions.

References

Akechi, H., Senju, A., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2010). The effect of gaze direction on the processing of facial expressions in children with autism spectrum disorder: An ERP study. *Neuropsychologia*, *48*(10), 2841–2851. http://doi.org/10.1016/j.neuropsychologia.2010.05.026

Amaral, D. G., Schumann, C. M., & Nordahl, C. W. (2008). Neuroanatomy of autism. *Trends in Neurosciences*, *31*(3), 137–145. http://doi.org/10.1016/j.tins.2007.12.005

American Clinical Neurophysiology Society. (2006). Guideline 5: Guidelines for Standard Electrode Position Nomenclature. *Clinical Neurophysiology*, *6*, 1–3. http://doi.org/10.1080/1086508X.2006.11079580

American Psychiatric Association. (2013a). *Diagnostic and Statistical Manual of Mental Disorders*. *American Psychiatric Publishing*. http://doi.org/10.1176/appi.books.9780890425596.744053

American Psychiatric Association. (2013b). Highlights of Changes from DSM-IV-TR to DSM-5. *American Psychiatric Publishing*, 19. http://doi.org/10.1176/appi.focus.11.4.525

Antoniadou, I., Manson, G., Dervilis, N., Barszcz, T., Staszewski, W. J., & Worden, K. (2012). Use of the Teager-Kaiser energy operator for condition monitoring of a wind turbine gearbox. In *Proceedings of the International Conference on Noise and Vibration Engineering ISMA 2012* (pp. 4255–4268).

Apicella, F., Sicca, F., Federico, R. R., Campatelli, G., & Muratori, F. (2013). Fusiform Gyrus responses to neutral and emotional faces in children with Autism Spectrum Disorders: A High Density ERP study. *Behavioural Brain Research*, *251*, 155–162. http://doi.org/10.1016/j.bbr.2012.10.040

Association for Applied Psychophysiology and Biofeedback. (2011). About Biofeedback. Retrieved August 25, 2016, from http://www.aapb.org/i4a/pages/index.cfm?pageid=3463

Atkinson, A. J. Jr., Colburn, W. A., DeGruttola, V. G., DeMets, D. L., Downing, G. J., Hoth, D. F., Oates, J. A., Peck, C. C., Schooley, R. T., Spilker, B. A., Woodcock, J., Zeger, S. L. (2001). Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clinical Pharmacology and Therapeutics*, *69*(3), 89–95. http://doi.org/10.1067/mcp.2001.113989

Batty, M., Meaux, E., Wittemeyer, K., Rogé, B., & Taylor, M. J. (2011). Early processing of emotional faces in children with autism: An event-related potential study. *Journal of Experimental Child Psychology*, *109*(4), 430–444. http://doi.org/10.1016/j.jecp.2011.02.001

Bayless, S. J., Glover, M., Taylor, M. J., & Itier, R. J. (2011). Is it in the eyes? Dissociating the role of emotion and perceptual features of emotionally expressive faces in modulating orienting to eye gaze. *Visual Cognition*, *19*(4), 483–510. http://doi.org/10.1080/13506285.2011.552895

Berggren, S., Engström, A-C., & Bölte, S. (2016). Facial affect recognition in autism, ADHD and typical development. *Cognitive Neuropsychiatry*, *6805*(April), 1–15. http://doi.org/10.1080/13546805.2016.1171205

Bernier, R., Aaronson, B., & McPartland, J. (2013). The role of imitation in the observed heterogeneity in EEG mu rhythm in autism and typical development. *Brain and Cognition*, *82*(1), 69–75. http://doi.org/10.1016/j.bandc.2013.02.008

Bernstein, M., & Yovel, G. (2015). Two neural pathways of face processing: A critical evaluation of current models. *Neuroscience and Biobehavioral Reviews*, *55*(JUNE 2015), 536–546. http://doi.org/10.1016/j.neubiorev.2015.06.010

Bluschke, A., Broschwitz, F., Kohl, S., Roessner, V., & Beste, C. (2016). The neuronal mechanisms underlying improvement of impulsivity in ADHD by theta/beta neurofeedback. *Scientific Reports*, *12*(6). http://doi.org/10.1038/srep31178

Bosl, W., Tierney, A., Tager-Flusberg, H., & Nelson, C. (2011). EEG complexity as a biomarker for autism spectrum disorder risk. *BMC Medicine*, *9*(1), 18. http://doi.org/10.1186/1741-7015-9-18

Brain Products GmbH. (2011). 64Ch Standard Electrode Layout for actiCHamp. Retrieved August 10, 2016, from http://actichamp.com/materials/actiCHamp_64Ch_Standard.pdf

Brain Products GmbH. (2016). actiCAP. Retrieved August 8, 2016, from http://www.brainproducts.com/productdetails.php?id=4

Cecotti, H., & Gräser, A. (2011). Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *33*(3), 433–445. http://doi.org/10.1109/TPAMI.2010.125

Cencini, M., Cecconi, F., & Vulpiani, A. (2010). *Chaos From Simple Models to Complex Systems*. Singapore: World Scientific.

Centers for Disease Control and Prevention. (2016). Autism Spectrum Disorder (ASD). Retrieved May 23, 2016, from http://www.cdc.gov/ncbddd/autism/data.html

Cheon, E.-J., Koo, B.-H., & Choi, J.-H. (2016). The Efficacy of Neurofeedback in Patients with Major Depressive Disorder: An Open Labeled Prospective Study. *Joulnal of Applied Psychophysiology and Biofeedback*, *41*(1), 103–110. http://doi.org/10.1007/s10484-015-9315-8

Clark, J. W. (2009). The Origin of Biopotentials. In J. G. Webster (Ed.), *Medical Instrumentation: Application and Design* (4th ed., pp. 126–188). Willey Global Education.

Coben, R., Linden, M., & Myers, T. E. (2010). Neurofeedback for autistic spectrum disorder: A review of the literature. *Applied Psychophysiology Biofeedback*, *35*(1), 83–105. http://doi.org/10.1007/s10484-009-9117-y

Coles, M. G. H., & Rugg, M. D. (1996). Event-related brain potentials: an introduction. In M. G. H. Coles & M. D. Rugg (Eds.), *Electrophysiology of Mind: Event-related Brain Potentials and Cognition* (pp. 1–26). Oxford University Press. http://doi.org/10.1093/acprof:oso/9780198524168.003.0001

Collura, T. F. (1997). EEG frequency bands. Retrieved July 14, 2016, from http://www.brainmaster.com/generalinfo/eegbands/eegbands.html

Crider, A., & Pillai, A. (2016). The Neurobiological Basis for Social Affiliation in Autism Spectrum Disorder and Schizophrenia. *Current Behavioral Neuroscience Reports*. http://doi.org/10.1007/s40473-016-0079-0

Davies, P. L., & Gavin, W. J. (2007). Validating the diagnosis of sensory processing disorders using EEG technology. *The American Journal of Occupational Therapy. : Official Publication of the American Occupational*

Therapy Association, *61*(2), 176–189.

Dawson, G., Webb, S. J., Carver, L., Panagiotides, H., & McPartland, J. (2004). Young children with autism show atypical brain responses to fearful versus neutral facial expressions of emotion. *Developmental Science*, *7*(3), 340–359. http://doi.org/10.1111/j.1467-7687.2004.00352.x

Dawson, G., Webb, S. Ja., & McPartland, J. (2005). Understanding the Nature of Face Processing Impairment in Autism: Insights From Behavioral and Electrophysiological Studies. *Developmental Neuropsychology*, *27*(3), 425–458. http://doi.org/10.1207/s15326942dn2703

De Jong, M. C., van England, H., & Kemner, C. (2008). Attentional Effects of Gaze Shifts Are Influenced by Emotion and Spatial Frequency, but Not in Autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*(4), 443–454. http://doi.org/10.1097/CHI.0b013e31816429a6

Deilami, M., Jahandideh, A., Kazemnejad, Y., Fakour, Y., Alipoor, S., Rabiee, F., Pournesaie, G. S., Noot Heidari, R. Mosavi, S. A. (2016). Case Report: The Effect of Neurofeedback Therapy on Reducing Symptoms Associated with Attention Deficit Hyperactivity Disorder: A Case Series Study. *Basic and Clinical Neuroscience Journal*, *7*(2), 167–171. http://doi.org/10.15412/J.BCN.03070211

Destexhe, A., & Bedard, C. (2012). Do neurons generate monopolar current sources? *Journal of Neurophysiology*, *108*, 953–955. http://doi.org/10.1152/jn.00357.2012

Dichter, G. S. (2012). Functional magnetic resonance imaging of autism spectrum disorders. *Dialogues in Clinical Neuroscience*, *14*(3), 319–351. http://doi.org/10.1016/j.brainres.2010.11.076.Structural

Duffy, F. H., & Als, H. (2012). A stable pattern of EEG spectral coherence distinguishes children with autism from neuro-typical controls - a large case control study. *BMC Medicine*, *10*(1), 64. http://doi.org/10.1186/1741-7015-10-64

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., Polich, J., Reinvang, I., Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, *120*(11), 1883–1908. http://doi.org/10.1016/j.clinph.2009.07.045

Eack, S., Mazefsky, C. A., & Minshew, N. (2013). Misinterpretation of Facial Expressions of Emotion in Verbal Adults with Autism Spectrum Disorder. *Autism*, *19*(3), 308–315. http://doi.org/10.1177/1362361314520755.Misinterpretation

Erfanian, A., Oveisi, F., & Shadvar, A. (2011). Feature Extraction by Mutual Information Based on Minimal-Redundancy-Maximal-Relevance Criterion and Its Application to Classifying EEG Signal for Brain-Computer Interfaces. In R. Fazel-Rezai (Ed.), *Recent Advances in Brain-Computer Interface Systems* (pp. 67–82). InTech. http://doi.org/10.5772/579

Fabiani, G. E., McFarland, D. J., Wolpaw, J. R., & Pfurtscheller, G. (2004). Conversion of EEG activity into cursor movement by a brain-computer interface (BCI). *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *12*(3), 331–338. http://doi.org/10.1109/TNSRE.2004.834627

Fabiani, M., Gratton, G., & Federmeier, K. D. (2007). Event-Related Brain Potentials: Methods, Theory, and Applications. In J. T. Cacioppo, L. G. Tassinary, & G. Berntson (Eds.), *Handbook of Psychophysiology* (3rd ed., pp. 85–119). Cambridge University Press. http://doi.org/http://dx.doi.org/10.1017/CBO9780511546396.004

Fan, J., Wade, J. W., Bian, D., Key, A. P., Warren, Z. E., Mion, L. C., & Sarkar, N. (2015). A Step Towards EEG-based Brain Computer Interface for Autism Intervention. In *Anual International Conference of the IEEE Engeneering in Medicine and Biology Society*. http://doi.org/10.1109/EMBC.2015.7319213

Faust, O., Acharya U, R., Krishnan, S. M., & Min, L. C. (2004). Analysis of cardiac signals using spatial filling index and time-frequency domain. *BioMedical Engineering Online*, *3*, 1–11. http://doi.org/10.1186/1475-925X-3-30

Fernández, T., Bosch-Bayard, J., Harmony, T., Caballero, M. I., Díaz-Comas, L., Galán, L., Ricardo-Garcell, J., Aubert, E., Otero-Ojeda, G. (2016). Neurofeedback in Learning Disabled Children: Visual versus Auditory Reinforcement. *Applied Psychophysiology Biofeedback*, *41*(1), 27–37. http://doi.org/10.1007/s10484-015-9309-6

Feuerriegel, D., Churches, O., Hofmann, J., & Keage, H. A. D. (2015). The N170 and face perception in psychiatric and neurological disorders: A systematic review. *Clinical Neurophysiology*, *126*(6), 1141–1158.

http://doi.org/10.1016/j.clinph.2014.09.015

Figueiredo, N., Silva, F., Georgieva, P., & Tomé, A. (2011). Advances in Non-Invasive Brain-Computer Interfaces for Control and Biometry. In R. Fazel-Rezai (Ed.), *Recent Advances in Brain-Computer Interface Systems* (pp. 171–192). InTech. http://doi.org/10.5772/579

Foss-Feig, J. H., McGugin, R. W., Gauthier, I., Mash, L. E., Ventola, P., & Cascio, C. J. (2016). A functional neuroimaging study of fusiform response to restricted interests in children and adolescents with autism spectrum disorder. *Journal of Neurodevelopmental Disorders*, *8*(1), 15. http://doi.org/10.1186/s11689-016-9149-6

Friedrich, E. V. C., Sivanathan, A., Lim, T., Suttie, N., Louchart, S., Pillen, S., & Pineda, J. A. (2015). An Effective Neurofeedback Intervention to Improve Social Interactions in Children with Autism Spectrum Disorder. http://doi.org/10.1007/s10803-015-2523-5

Friedrich, E. V. C., Suttie, N., Sivanathan, A., Lim, T., Louchart, S., & Pineda, J. a. (2014). Brain-computer interface game applications for combined neurofeedback and biofeedback treatment for children on the autism spectrum. *Frontiers in Neuroengineering*, *7*(July), 1–7. http://doi.org/10.3389/fneng.2014.00021

Frith, U., Happé, F., & Happe, F. (2005). Autism spectrum disorder. *Current Biology*, *15*(19), R786–90. http://doi.org/10.1016/j.cub.2005.09.033

Garman, H. D., Spaulding, C. J., Webb, S. J., Mikami, A. Y., Morris, J. P., & Lerner, M. D. (2016). Wanting it Too Much: An Inverse Relation Between Social Motivation and Facial Emotion Recognition in Autism Spectrum Disorder. *Child Psychiatry and Human Development*, 1–13. http://doi.org/10.1007/s10578-015-0620-5

Giorgio. (2016). Feature Selection for Matlab. MathWorks. Retrieved May, 2016 from https://www.mathworks.com/matlabcentral/fileexchange/54764-feature-selection-for-matlab

Gross, E., El-Baz, A. S., Sokhadze, G. E., Sears, L., Casanova, M. F., & Sokhadze, E. M. (2012). Induced EEG Gamma Oscillation Alignment Improves Differentiation Between Autism and Adhd Group Responses in a Facial Categorization Task. *Journal of Neurotherapy*, *16*(2), 78–91. http://doi.org/10.1080/10874208.2012.677631

Gu, Q., Li, Z., & Han, J. (2012). *Generalized Fisher Score for Feature Selection*. *Computer Science - Learning*. Retrieved May, 2016 from http://arxiv.org/abs/1202.3725\nhttp://researchbank.rmit.edu.au/view/rmit:160103

Guger, C., Edlinger, G., & Krausz, G. (2011). Hardware / Software Components and Applications of BCIs. In R. Fazel-Rezai (Ed.), *Recent Advances in Brain-Computer Interface Systems* (pp. 1–24). InTech. http://doi.org/10.5772/579

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial Emotion Recognition in Autism Spectrum Disorders: A Review of Behavioral and Neuroimaging Studies. *Neuropsychology Review*, *20*(3), 290–322. http://doi.org/10.1007/s11065-010-9138-6

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. http://doi.org/10.1016/S1364-6613(00)01482-0

Hendrika, P., Vlamings, J. M., Jonkman, L. M., Van Daalen, E., Van Der Gaag, R. J., & Kemner, C. (2010). Basic abnormalities in visual processing affect face processing at an early age in autism spectrum disorder. *Biological Psychiatry*, *68*(12), 1107–1113. http://doi.org/10.1016/j.biopsych.2010.06.024

Heunis, T., Aldrich, C., & de Vries, P. J. (2016). Recent advances in resting state electroencephalography biomarkers for autism spectrum disorder – a review of methodological and clinical challenges. *Pediatric Neurology*, *61*, 28–37. http://doi.org/10.1016/j.pediatrneurol.2016.03.010

Jamal, W., Das, S., Oprescu, I.-A., Maharatna, K., Apicella, F., & Sicca, F. (2014). Classification of autism spectrum disorder using supervised learning of brain connectivity measures extracted from synchrostates. *Journal of Neural Engineering*, *11*(4), 046019. http://doi.org/10.1088/1741-2560/11/4/046019

Jeste, S. S., Frohlich, J., & Loo, S. K. (2015). Electrophysiological biomarkers of diagnosis and outcome in neurodevelopmental disorders. *Current Opinion in Neurology*, *28*(2), 110–116. http://doi.org/10.1097/WCO.0000000000000181

Johan. (2007). The problem with comparing faces to other stimuli. Retrieved July 31, 2016, from https://phineasgage.wordpress.com/2007/05/05/the-problem-with-comparing-faces-to-other-stimuli/

Jurcak, V., Tsuzuki, D., & Dan, I. (2007). 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems. *NeuroImage*, *34*(4), 1600–1611. http://doi.org/10.1016/j.neuroimage.2006.09.024

Kanner, L. (1943). Autistic Disturbances of Affective Contact. *Nervous Child*, *2*, 217–250.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *361*(1476), 2109–28. http://doi.org/10.1098/rstb.2006.1934

Kashihara, K. (2014). A Brain-Computer Interface for Potential Nonverbal Facial Communication Based on EEG Signals Related to Specific Emotions. *Frontiers in Neuroscience*, *8*(8 JUL), 1–12. http://doi.org/10.3389/fnins.2014.00244

Kliková, B., & Raidl, A. (2011). Reconstruction of Phase Space of Dynamical Systems Using Method of Time Delay. In *Proceedings of the 20th Annual Conference of Doctoral Students - WDS 2011* (pp. 83–87). Prague, Czech Republic.

Koles, Z. J., Lazar, M. S., & Zhou, S. Z. (1990). Spatial patterns underlying population differences in the background EEG. *Brain Topography*, *2*(4), 275–284. http://doi.org/10.1007/BF01129656

Korkmaz, B. (2011). Theory of mind and neurodevelopmental disorders of childhood. *Pediatric Research*, *69*(5 PART 2). http://doi.org/10.1203/PDR.0b013e318212c177

Kouijzer, M. E. J., van Schie, H. T., Gerrits, B. J. L., Buitelaar, J. K., & de Moor, J. M. H. (2013). Is EEG-biofeedback an Effective Treatment in Autism Spectrum Disorders? A Randomized Controlled Trial. *Applied Psychophysiology and Biofeedback*, *38*(1), 17–28. http://doi.org/10.1007/s10484-012-9204-3

Lai, M.-C., Lombardo, M. V, & Baron-Cohen, S. (2014). Autism. *The Lancet*, *383*(9920), 896–910. http://doi.org/10.1016/S0140-6736(13)61539-1

Lerner, M. D., McPartland, J. C., & Morris, J. P. (2013). Multimodal emotion processing in autism spectrum disorders: An event-related potential study. *Developmental Cognitive Neuroscience*, *3*, 11–21.

http://doi.org/10.1016/j.dcn.2012.08.005

Luck, S. J. (2005). An Introduction to Event-Related Potentials and Their Neural Origins. In S. J. Luck (Ed.), *An Introduction to the Event-Related Potential Technique*. MIT Press.

Luckhardt, C., Jarczok, T. A., & Bender, S. (2014). Elucidating the neurophysiological underpinnings of autism spectrum disorder: New developments. *Journal of Neural Transmission*, *121*(9), 1129–1144. http://doi.org/10.1007/s00702-014-1265-4

Ludwig, K. A., Miriani, R. M., Langhals, N. B., Joseph, M. D., Anderson, D. J., & Kipke, D. R. (2009). Using a Common Average Reference to Improve Cortical Neuron Recordings From Microelectrode Arrays. *Journal of Neurophysiology*, *101*(3), 1679–1689. http://doi.org/10.1152/jn.90989.2008

Martínez, J., & Barrientos, A. (2011). Applied Advanced Classifiers for Brain Computer Interface. In R. Fazel-Rezai (Ed.), *Recent Advances in Brain-Computer Interface Systems* (pp. 25–66). InTech. http://doi.org/10.5772/579

Marzbani, H., Marateb, H. R., & Mansourian, M. (2016). Neurofeedback: A Comprehensive Review on System Design, Methodology and Clinical Applications. *Basic and Clinical Neuroscience Journal*, *7*(2), 143–158. http://doi.org/10.15412/J.BCN.03070208

Mathworks. (2016). Measure the Power of a Signal. Retrieved August 13, 2016, from http://www.mathworks.com/help/signal/ug/measure-the-power-of-a-signal.html

Matlis, S., Boric, K., Chu, C. J., & Kramer, M. A. (2015). Robust disruptions in electroencephalogram cortical oscillations and large-scale functional networks in autism. *BMC Neurology*, *15*(1), 97. http://doi.org/10.1186/s12883-015-0355-8

Maxwell, C. R., Villalobos, M. E., Schultz, R. T., Herpertz-Dahlmann, B., Konrad, K., & Kohls, G. (2015). Atypical Laterality of Resting Gamma Oscillations in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *45*, 292–297. http://doi.org/10.1007/s10803-013-1842-7

Medscape. (2014). Normal EEG Waveforms. Retrieved July 14, 2016, from http://emedicine.medscape.com/article/1139332-overview#a1

Nazari, M. A. (2012). EEG findings in ADHD and the application of EEG biofeedback in treatment of ADHD. In J. M. Norvilitis (Ed.), *Current directions*

in ADHD and its treatment (pp. 269–286). InTech. http://doi.org/10.5772/1213

Nazari, M. A., Mosanezhad, E., Hashemi, T., & Jahan, A. (2012). The effectiveness of neurofeedback training on EEG coherence and neuropsychological functions in children with reading disability. *Clinical EEG and Neuroscience*, *43*(4), 315–322. http://doi.org/10.1177/1550059412451880

Nunez, P. L., & Srinivasan, R. (2006). The Physics-EEG Inferface. In P. L. Nunez & R. Srinivasan (Eds.), *Electric Fields of the Brain: The Neurophysics of EEG* (2nd ed.). Oxford University Press.

O'Connor, K., Hamm, J. P., & Kirk, I. J. (2005). The neurophysiological correlates of face processing in adults and children with Asperger's syndrome. *Brain and Cognition*, *59*, 82–95. http://doi.org/10.1016/j.bandc.2005.05.004

Ojala, M., & Garriga, G. C. (2009). Permutation tests for studying classifier performance. In *Proceedings - IEEE International Conference on Data Mining, ICDM* (pp. 908–913). http://doi.org/10.1109/ICDM.2009.108

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, *112*(4), 713–719. http://doi.org/http://dx.doi.org/10.1016/S1388-2457(00)00527-7

Otten, L. J., & Rugg, M. D. (2005). Interpreting Event-Related Brain Potentials. In T. C. Handy (Ed.), *Event-related Potentials: A Methods Handbook* (pp. 3–16). Press, MIT.

Pfurtscheller, G., & Lopes, F. H. (1999). Event-related EEG / MEG synchronization and desynchronization : basic principles. *Clinical Neurophysiology*, *110*, 1842–1857.

Pincus, S. M., Gladstone, I. M., & Ehrenkranz, R. A. (1991). A REGULARITY STATISTIC FOR MEDICAL DATA ANALYSIS. *Journal of Clinical Monitoring and Computing*, *7*(4), 335–345. http://doi.org/10.1007/BF01619355

Pistorius, T., Aldrich, C., Auret, L., & Pineda, J. (2013). Early detection of risk of autism spectrum disorder based on recurrence quantification analysis of electroencephalographic signals. In *6th Annual International IEEE EMBS Conference on Neural Engineering*. http://doi.org/10.1109/NER.2013.6695906

Ploog, B. O., Scharf, A., Nelson, D., & 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*, 301–322. http://doi.org/10.1007/s10803-012-1571-3

Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology, Heart and Circulatory Physiology*, *278*(6), 2039–2049.

Schabus, M., Heib, D. P. J., Lechinger, J., Griessenberger, H., Klimesch, W., Pawlizki, A., … Hoedlmoser, K. (2014). Enhancing sleep quality and memory in insomnia using instrumental sensorimotor rhythm conditioning. *Biomoligal Psychology*, *95*, 126–134. http://doi.org/10.1016/j.biopsycho.2013.02.020

Simões, M., Lima, J., Direito, B., Castelhano, J., Ferreira, C., Carvalho, P., & Castelo-branco, M. (2015). Feature analysis for correlation studies of simultaneous EEG-fMRI data : a proof of concept for neurofeedback approaches. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 4065–4068). http://doi.org/10.1109/EMBC.2015.7319287

Strehl, U., Birkle, S. M., Wörz, S., & Kotchoubey, B. (2014). Sustained Reduction of Seizures in Patients with Intractable Epilepsy after Self-Regulation Training of Slow Cortical Potentials - 10 Years After. *Frontiers in Human Neuroscience*, *8*(604), 1–7. http://doi.org/10.3389/fnhum.2014.00604

Surmeli, T., Ertem, A., Eralp, E., & Kos, I. H. (2012). Schizophrenia and the efficacy of qEEG-guided neurofeedback treatment: a clinical case series. *Clinical EEG and Neuroscience*, *43*(2), 133–144. http://doi.org/10.1177/1550059411429531

Tatum, W. O. (2014). Normal EEG. In W. O. Tatum (Ed.), *Handbook of EEG Interpretation* (2nd ed., pp. 1–56). Demos Medical Publishing.

Tierney, A. L., Gabard-Durnam, L., Vogel-Farley, V., Tager-Flusberg, H., & Nelson, C. A. (2012). Developmental trajectories of resting eeg power: An endophenotype of autism spectrum disorder. *PLoS ONE*, *7*(6). http://doi.org/10.1371/journal.pone.0039127

Trans Cranial Technologies ldt. (2012). 10 / 20 System Positioning Manual. *Trans Cranial Technologies*. Hong Kong: Trans Cranial Technologies ldt. Retrieved

from http://www.trans-cranial.com/local/manuals/10_20_pos_man_v1_0_pdf.pdf

Tye, C., Battaglia, M., Bertoletti, E., Ashwood, K. L., Azadi, B., Asherson, P., … McLoughlin, G. (2014). Altered neurophysiological responses to emotional faces discriminate children with ASD, ADHD and ASD+ADHD. *Biological Psychology*, *103*, 125–34. http://doi.org/10.1016/j.biopsycho.2014.08.013

Ulrich, T. J. (2006). *Envelope calculation from the Hilbert transform*.

Wagner, J. B., Hirsch, S. B., Vogel-Farley, V. K., Redcay, E., & Nelson, C. A. (2013). Eye-Tracking, Autonomic and Electrophysiological Correlates of Emotional Face Processing in Adolescents with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *43*(1), 188–199. http://doi.org/10.1007/s10803-012-1565-1.

Wainer, A. L., & Ingersoll, B. R. (2011). The use of innovative computer technology for teaching social communication to individuals with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *5*(1), 96–107. http://doi.org/10.1016/j.rasd.2010.08.002

Walsh, J. A., Creighton, S. E., & Rutherford, M. D. (2016). Emotion Perception or Social Cognitive Complexity : What Drives Face Processing Deficits in Autism Spectrum Disorder ? *Journal of Autism and Developmental Disorders*, *46*(2), 615–623. http://doi.org/10.1007/s10803-015-2606-3

Walsh, P., Elsabbagh, M., Bolton, P., & Singh, I. (2011). In search of biomarkers for autism: scientific, social and ethical challenges. *Nature Reviews Neuroscience*, *12*(10), 603–612. http://doi.org/10.1038/nrn3113

Wang, Jun., Barstein, J., Ethridge, L. E., Mosconi, M. W., Takarae, Y., & Sweeney, J. a. (2013). Resting state EEG abnormalities in autism spectrum disorders. *Journal of Neurodevelopmental Disorders*, *5*(1), 1–14. http://doi.org/10.1186/1866-1955-5-24

Wang, Jue., Yan, N., Liu, H., Liu, M., & Tai, C. (2007). Brain-Computer Interfaces Based on Attention and Complex Mental Tasks. *Digital Human Modeling*, *4561*, 467–473. http://doi.org/10.1007/978-3-540-73321-8

Wang, Q., Sourina, O., & Nguyen, M. K. (2010). EEG-based “serious” games design for medical applications. *Proceedings - 2010 International Conference on Cyberworlds, CW 2010*, 270–276. http://doi.org/10.1109/CW.2010.56

Wong, T. K. W., Fung, P. C. W., Chua, S. E., & McAlonan, G. M. (2008). Abnormal spatiotemporal processing of emotional facial expressions in childhood autism: dipole source analysis of event-related potentials. *European Journal of Neuroscience*, *28*(2), 407–416. http://doi.org/10.1111/j.1460-9568.2008.06328.x

Yoshimura, N., & Itakura, N. (2011). Usability of Transient VEPs in BCIs. In R. Fazel-Rezai (Ed.), *Recent Advances in Brain-Computer Interface Systems* (pp. 119–134). InTech. http://doi.org/10.1109/CNE.2007.369683