POLITECNICO DI TORINO

MASTER'S DEGREE IN BIOMEDICAL ENGINEERING

MASTER'S DEGREE THESIS

EARLY DETECTION OF ALZHEIMER'S DISEASE IN aMCI PATIENTS: EYE-TRACKING ANALYSIS USING NAP4A PLATFORM

Supervisors

Prof. Filippo MOLINARI

Co-Advisors

Ing. Alberto SANNA

Ing. Alessandra GARGIULO

Ing. Alice LEOPORINI

Candidate

Simona MUSICCO

Academic Year 2024/2025

Abstract

Dementias represent a broad group of neurodegenerative disorders characterized by progressive cognitive decline. Alzheimer's disease (AD) is a degenerative neurological condition responsible for 60-80% of cases of irreversible cognitive impairment. Amnestic mild cognitive impairment (aMCI) is a condition marked by reduced cognitive abilities and is considered a possible transitional state between normal aging and AD. Therefore, detecting the early stages of AD is a challenge for traditional methodologies such as brain imaging and qualitative tests. Given the significant impact of AD on patients and their families, early identification is crucial to optimize therapeutic strategies and slow the progression of the disease. The NAP4A2020 study (Neurocognitive Assessment Platform 4 Alzheimer) follows a low-risk, longitudinal protocol aimed at identifying the preclinical phase of Alzheimer's disease (AD) and estimating the probability of progression from amnestic Mild Cognitive Impairment (aMCI) to AD. All participants underwent a standard neuropsychological assessment to evaluate different cognitive domains (memory, language, processing speed) and a battery of digital cognitive tests (CANTAB). Additionally, they were exposed to two sets of visual stimuli: personal images and images referred to as "Neffie", selected for their emotional and attentional impact. The NAP4A platform, developed by the Center for Advanced Technology in Health and Wellbeing in collaboration with the Neurology Unit of San Raffaele Hospital, integrates multiple sensors for acquiring neurophysiological signals, including galvanic response, electroencephalographic activity, photoplethysmographic data, eye-tracking, and facial expression analysis. This thesis project focused on analyzing eye-tracking (TOBII™, desktop-based, binocular sampling frequency of 90 Hz) as a digital tool for cognitive and behavioral assessment in 24 participants (3 with aMCI and 21 healthy controls). Specifically, the study aimed to evaluate the presence of differences between aMCI participants and healthy controls based on the extracted ocular parameters. In healthy controls, the results showed a higher number of stable fixations, greater saccadic accuracy, and lower spatial dispersion of fixations, indicative of efficient visual processing and proper attention management. Preliminary findings suggest that aMCI patients exhibit differences compared to healthy controls, characterized by a lower number and shorter duration of fixations, higher saccadic frequency, indicative of fragmented and less efficient visual exploration, and reduced visual entropy. This last parameter suggests alterations in cognitive processes related to attention and memory. The analysis of neuropsychological tests confirmed a decline in performance in aMCI patients compared to healthy controls,

particularly in tests assessing episodic memory, sustained attention, and executive functions. Preliminary sample analysis suggests that eye-tracking, combined with neuropsychological testing, could serve as a sensitive biomarker for the early monitoring of cognitive decline, improving the early diagnosis of Alzheimer's disease and enabling a more precise and continuous evaluation of disease progression.

Abstract in Lingua Italiana

Le demenze rappresentano un ampio gruppo di patologie neurodegenerative caratterizzate da un deterioramento progressivo delle funzioni cognitive. L'Alzheimer (AD) è una patologia neurologica degenerativa, responsabile del 60-80% dei casi di declino cognitivo irreversibile. Il deterioramento cognitivo amnestico (aMCI) è una condizione caratterizzata da una riduzione delle capacità cognitive, considerata un possibile stato di transizione tra l'invecchiamento fisiologico e l'AD. Pertanto, il rilevamento delle fasi iniziali di AD è una sfida per le metodologie tradizionali quali esami di imaging cerebrale e test qualitativi. Dato l'impatto significativo dell'AD sulla vita dei pazienti e dei loro familiari, l'identificazione precoce è fondamentale per ottimizzare le strategie terapeutiche e rallentare la progressione della malattia. Lo studio NAP4A2020, Neurocognitive Assessment Platform 4 Alzheimer, prevede un protocollo a basso rischio, longitudinale, finalizzato all'identificazione della fase preclinica di AD, indicando il grado di probabilità di evoluzione da aMCI ad AD. Tutti i partecipanti sono stati sottoposti ad una valutazione neuropsicologica standard, volta ad indagare diversi domini cognitivi (memoria, linguaggio, velocità di elaborazione) e una batteria di test cognitivi digitali (CANTAB). Inoltre, i partecipanti sono stati esposti a due serie di stimoli visivi: immagini personali e immagini definite "Neffie" scelte per il loro impatto emotivo ed attentivo. La piattaforma NAP4A, sviluppata dal Center for Advanced Technology in Health and Wellbeing in collaborazione con l'Unità di Neurologia dell'Ospedale San Raffaele, integra diversi sensori per l'acquisizione dei segnali quali risposta galvanica, elettroencefalografica, fotopletismografico, eye-tracking ed espressione facciale. Il presente progetto di tesi si è concentrato sull'analisi dell'eye-tracker (TOBII™, desktop-based, frequenza di campionamento binoculare di 90 Hz) come strumento digitale di valutazione cognitiva e comportamentale in 24 partecipanti (3 aMCI e 21 controlli sani). Nello specifico, si è andata a valutare la presenza di differenze tra partecipanti aMCI e controlli sani evidenziate nei parametri oculari estratti. Nei soggetti di controllo, i risultati hanno mostrato un numero maggiore di fissazioni stabili, una maggiore precisione delle saccadi e una minore dispersione spaziale delle fissazioni, indicative di un'elaborazione visiva efficiente e di un'adeguata gestione dell'attenzione. I risultati preliminari indicano che i pazienti con aMCI mostrano differenze rispetto ai controlli sani, caratterizzate da un minor numero delle fissazioni e minore durata, una maggiore frequenza delle saccadi, indice di un'esplorazione visiva frammentata e meno efficiente, e una riduzione dell'entropia visiva. Quest'ultimo dato suggerisce alterazioni nei processi cognitivi legati all'attenzione e alla memoria. L'analisi dei

test neuropsicologici ha confermato una riduzione delle prestazioni nei pazienti con aMCI rispetto ai controlli sani, in particolare nei test che valutano memoria episodica, attenzione sostenuta e funzioni esecutive. Dall'analisi preliminare del campione i risultati ottenuti suggeriscono che l'eye-tracking, combinato con test neuropsicologici, potrebbe rappresentare un biomarcatore sensibile per il monitoraggio precoce dei segni di deterioramento cognitivo migliorando la diagnosi precoce della malattia di Alzheimer e consentendo una valutazione più precisa e continua del decorso della malattia.

A Francesco,
per avermi fatto vedere il colore
quando tutto sembrava essere in bianco e nero

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Neurodegenerative Disease

According to the World Health Organization, the term dementia refers to numerous disorders of an organic nature of the brain, characterized by the global and progressive impairment of previously acquired cognitive functions, specifically memory, logical-deductive reasoning, language, and the ability to orient topographically and temporally [1]. Dementia also significantly interferes with interpersonal relationships and a person's ability to perform activities of daily living, such as work, social ability. This definition was provided by the World Health Organization. Alzheimer's disease is the most common form of dementia and accounts for 60-70% of cases. The term dementia was first used in the medical field in the volume "De Medicina" by Aulus Cornelius Celsus (written in the 1st century AD) to refer to conditions of altered intelligence and behavior. In the 1900s, thanks to the development of neuropathology observation techniques, neuropsychiatrist Kraepelin outlined the first aspects of modern framing of dementia by introducing the concept of organic dementia, to define morbid conditions, due to diseases of the central nervous system, that are characterized by a precise neuropathological picture and irreversible loss of intellectual capacity. The increasing availability of techniques for studying the functioning of the central nervous system and the advancement of neuropathological techniques and knowledge have led to a greater clinical characterization of dementias and their distinction from both psychosis and the changes in cognitive function found with aging[2]. Nowadays, reference is made to the clinical criteria defined by the 2000 DSM-IV TR (Diagnostic and Statistical Manual of Mental disorders Text Revision) of the American Psychiatric Association and the 1922 ICD-10 (International Statistical Classification of Diseases and Related Health Problems) of the World Health Organization, which allow differentiation from other pathological conditions in which cognitive impairment can be found. Symptomatology presents cognitive symptoms and noncognitive symptoms inherent in the sphere of personality, affectivity, ideation and perception, and vegetative functions. The severity of changes in intellectual functions is such that it interferes with the daily acts of an individual's life and causes transformations in personality and mood. Dementias represent one of the main causes of disability in the world population with a significant socio-health and economic impact: an increasing number of people are affected, requiring a qualified network of services that integrates health and social care[3]. The aging population is mainly responsible for the gradual increase in the incidence of this disease.

1.1.1 Type of dementia

Dementias can be classified according to several criteria, for example, according to the age of onset (senile or presenile, depending on whether they arise before or after the sixth to seventh decade of life), the site of lesion (cortical and subcortical) or to prognosis (irreversible degenerative or reversible degenerative). The most commonly used criterion is the etiologic one, which divides dementias into two categories: idiopathic or degenerative primary dementias (for which the development of symptomatology is attributable to neuronal degeneration) and secondary dementias (for which cognitive impairment is caused by a defined pathology and are divided into vascular, hydrostatic and mechanostructural, transmissible, inflammatory, paraneoplastic, toxic, nutritional and deficiency, dysendocrine, dysmetabolic, and psychiatric). Idiopathic dementias are divided into Alzheimer's and non-Alzheimer's. The latter category includes fronto-temporal dementias, dementias with Lewy bodies, and movement disorders with dementia. Alzheimer's disease, the most common form of dementia, accounts for 60-70% of cases. Other major forms include vascular dementia, dementia with Lewy bodies, and a group of diseases that contribute to frontotemporal dementia. The boundaries between different forms of dementia are not well defined, and mixed forms often coexist. Secondary dementias may improve depending on treatment of the underlying pathology that caused them, while primary dementia is degenerative in nature so the process of cognitive impairment is developmental. Considering the various forms of dementia, Stefanovski et al. in [4] conducted a study of a group of people with an average age of 97.7 years evaluating the following distributions: 65% of dementia suffer from Alzheimer's disease, about 14% from vascular dementia or a combination of Alzheimer's and vascular dementia, 12% from Lewy body dementia, and finally the remaining 9 percent are attributed to dementia of various types: degenerative, metabolic, endocrine, toxic, and tumor. The figure shows information regarding the distribution of dementia types in the population.[5]

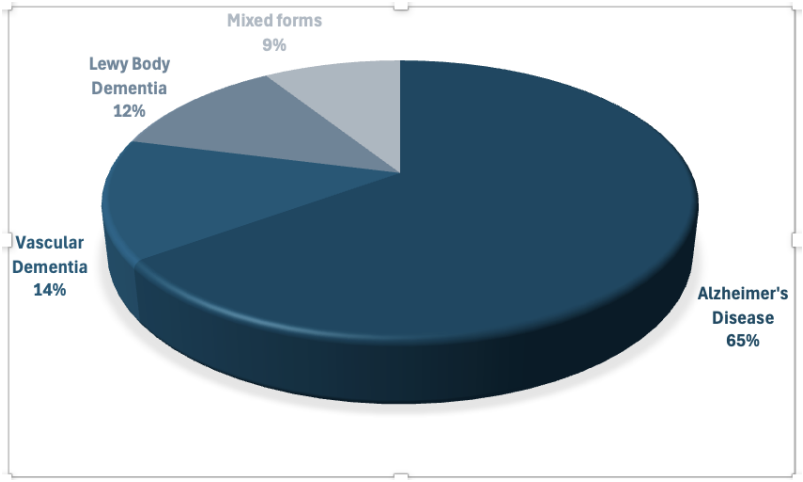


Figure 1.1: Epidemiology of different types of dementia.

According to to the World Organization, over 55 million people worlwide are currently living with dementia, with nearly 10 million new cases each year[1]. Projection indicate that the total number of cases could reach 139 million by 2050[6]. As for Europe and Italy, in 2023 the over-65 age group represented about 21.1% of the total population[7]. Italy, according to 2023 ISTAT data, holds the second position in Europe for life expectancy, at 83,1 years. Moreover, according to recent statistics from ISTAT 2023, 24.1% of the Italian population was over 65 years old. Globally, dementia is the fifth leading cause of death. In Italy, the number of people with dementia is estimated to be about 1.48 million, including about 600,000 diagnosed with Alzheimer's disease, according to data presented during the 18th National Congress of the Italian Society of Neurological Rehabilitation in 2018.

1.1.2 Diagnosis, symptoms and treatment of dementia

Dementia is a disease that involves a progressive deterioration of cognitive functions, such as memory, orientation, language, and reasoning skills. The recognition of this disease is done through clinical and neuropsychological criteria, and the use of neuroimaging (brain imaging techniques) is particularly useful in differentiating the various forms of dementia. During the evolution of a degenerative type of dementia, we can distinguish two main phases: the preclinical phase and the clinical phase. In the preclinical stage, which can

last for a long time, cognitive decline is still mild and does not significantly interfere with daily activities, such as working or taking care of oneself. At this stage, patients are able to maintain some level of independence through compensatory strategies that help them overcome cognitive difficulties. The ability to implement these strategies depends greatly on the individual's cognitive reserve, that is, the brain's ability to adapt to impairment and stress. People with higher cognitive reserve tend to show symptoms of dementia later, while those with lower cognitive reserve may develop them earlier. Over time, however, the situation worsens, and in the clinical stage, patients gradually lose autonomy in daily activities and become increasingly dependent on the help of a caregiver. The ability to perform simple tasks, such as grocery shopping, preparing meals, or managing finances, may fail, and this marks significant functional disability. To arrive at a definitive diagnosis of Alzheimer's, the most certain modality is the analysis of tissue samples, which can be done through biopsies taken from living people or, more commonly, by postmortem autopsy. The latter is preferred over other invasive techniques because there are no treatments that can stop or slow the progression of the disease, so post-mortem diagnosis is considered more accurate and scientifically useful[4]. In the advanced clinical stage of dementia, symptoms are greatly exacerbated. Cognitive difficulties that initially represented simply temporary difficulties turn into real incapacities. Increasingly severe amnestic deficits, spatial and temporal disorientation, and language disorders such as aphasia are manifested. In addition, other motor difficulties also appear, such as apraxia (difficulty in performing voluntary actions), alexia (inability to read) and agraphia (difficulty in writing), as well as deficits in abstract reasoning skills and in logic and judgment. Alongside these cognitive symptoms, there are also no-cognitive manifestations, which include changes in behavior and mood. Depression, irritability, and aggression are common, and in some cases strange or aggressive behavior may occur. These symptoms not only affect the person's daily life, but also result in severe limitations in his or her ability to perform social roles and participate in normal activities of daily living, such as working or maintaining relationships. Because dementia is a progressive disease, symptoms tend to worsen with time. Unfortunately, in advanced forms of dementia, such as Alzheimer's disease, there are no treatments that can stop or slow down the disease. However, there are medications that can temporarily improve some symptoms, such as memory and orientation problems. These drugs, unfortunately, do not stop cognitive impairment, but they do help maintain a higher level of independence for a period of time. In addition to medications, non-pharmacological therapies, such as cognitive stimulation, music therapy, and other forms of psychological and social support,

are also available that can alleviate some of the symptoms and improve the quality of life for patients and their families[8]. Although these therapies cannot stop the progression of the disease, they are still helpful in managing symptoms and improving patients' ability to interact and psychological well-being[9].

1.2 Alzheimer Disease

Alzheimer's disease is the most common form of dementia globally, accounting for about 65% of all dementia cases. It is characterized by a slow and progressive deterioration of cognitive, behavioral and functional functions. The condition appears to be more common in the female sex than in the male. Alzheimer's disease is classified among the primary cortical dementias and is described as a degenerative process that gradually destroys brain cells, increasingly impairing the individual's ability to lead a normal life[10]. The disease develops through three main stages:

- Mild phase: characterized by mild cognitive deficits, such as difficulty remembering recent events or finding the right words. In this stage, the person can still perform most daily activities independently, although mild difficulties may emerge that are often attributed to normal aging.
- Intermediate stage: symptoms become more pronounced and begin to interfere with daily activities. More pronounced problems with memory, confusion, difficulty recognizing familiar people and performing complex tasks occur. Changes in behavior, such as irritability or apathy, may appear.
- Advanced stage: Cognitive impairment is severe, and the person loses the ability to respond to the environment, converse, and possibly control movements. He or she requires full assistance with daily activities, and physical functions deteriorate, increasing the risk of infection and other complications.

1.2.1 Pathophysiology

Researchers in the field of neuropathology have identified the abnormal presence of amyloid plaques and aggregates of hyperphosphorylated tau proteins in the brains of Alzheimer's disease patients. These alterations are accompanied by a marked reduction in the number of neurons and dysfunction in neurotransmission processes. Based on these observations, it is hypothesized that the accumulation of these proteins may impair communication between nerve cells, progressively leading to their degeneration and death. In the context of Alzheimer's disease, two main categories of alterations can be identified, both of which are linked to the evolution of the disease and the onset of symptoms:

- Positive-type alterations, related to the formation of amyloid plaques, aggregates of tau proteins and other deposits in brain tissue.
- Alterations of the negative type, characterized by processes of atrophy due to the progressive loss of neurons and synaptic connections[11].

Analyses conducted on postmortem brain tissues of subjects with the disease confirmed the presence of neurofibrillary tangles with tau and accumulations of neuritic beta-amyloid (Aβ) plaques, consisting of sequences of 40 or 42 amino acids (Aβ40 and Aβ42). The synthesis of this protein is mediated by the transmembrane glycoprotein beta-APP (amyloid protein precursor or amyloid protein precursor) [4].

These deposits, distributed mainly in the hippocampus, amygdala and cerebral cortex, trigger inflammation that stimulates the immune system, leading to the intervention of cells such as macrophages and neutrophils. Such an immune response, however, results in irreversible damage to neurons. In addition, amyloid plaques alter the stability of the cell membrane, promoting abnormal calcium accumulation within cells and accelerating the process of neuronal death. However, some aspects of the pathogenesis of the disease remain unclear. In fact, it has been observed that beta-amyloid protein is also present in the brains of healthy elderly individuals, suggesting that the problem may lie in the mechanisms of elimination and disposal of the protein rather than in its simple formation. A crucial role in disease progression is also played by the Tau protein, which under normal conditions stabilizes axon structure through its interaction with microtubules. In individuals with dementia, this protein undergoes chemical alterations that result in its hyperphosphorylation and

subsequent aggregation into filamentous structures called Neurofibrillary Tangles (NFTs). These protein clusters impair microtubule function, promoting the loss of structural stability of the neuron and its subsequent deterioration. The most characteristic symptoms of Alzheimer's, such as memory impairment and speech difficulties, result from the inability of neurons to efficiently transmit nerve impulses. Progressive degeneration of nerve cells also results in thinning of gray matter, a distinctive phenomenon of the disease. Synaptic deficit, caused by the difficulty of communication between neurons, is the main culprit in the clinical manifestations of dementia[11].

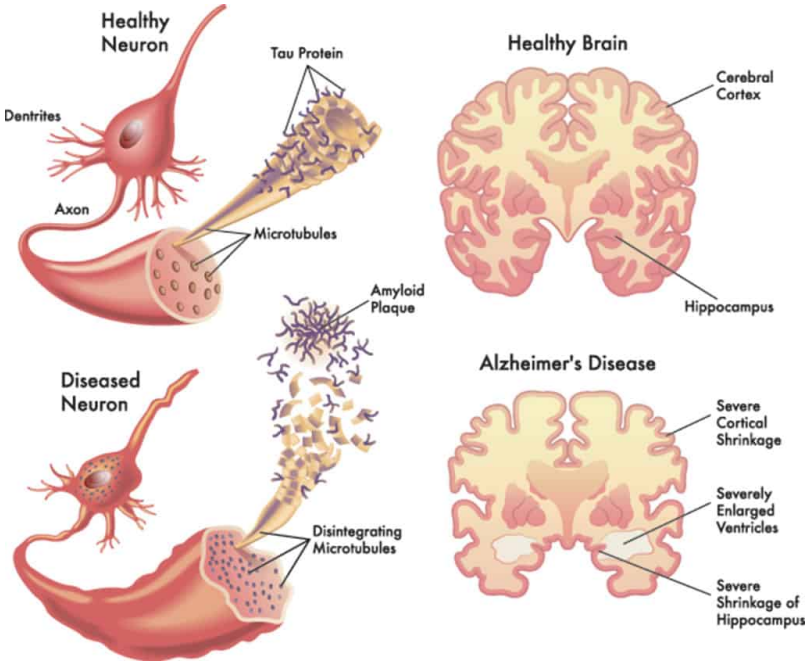


Figure 1.2: Differences between a healthy brain and a brain with Alzheimer's disease: the image shows the abnormal accumulation of tau and beta-amyloid proteins and the reduction of brain volume, which are characteristic of the disease.

1.2.2 Clinical Manifestations

It is very difficult to distinguish early signs of dementia, as they can occur sporadically even in healthy people, both young and old, although more frequently in individuals with dementia. The Alzheimer's Association of the USA has published a list of the top 10

premonitory symptoms of the onset of AD (common to all dementias) that may arouse suspicion and should be kept under observation:

1. Memory loss: forgetfulness of recently learned information such that daily life is impaired, reliance on others for activities previously done independently, increased need to rely on memory aids;
2. Difficulty planning or problem solving: problems following recipes, accounts or complex plans, difficulty concentrating;
3. Difficulty completing family commitments at home, work or leisure: difficulty completing daily tasks that result in problems driving the car to a familiar place, managing a budget at work or remembering the rules of a game;
4. Confusion with time or place: loss of sense of dates, seasons, and the passage of time;
5. Difficulty understanding visual images and spatial relationships: difficulty reading, judging distance, color or contrast;
6. New problems with words in speaking or writing: difficulty continuing a sentence during a conversation, high probability of repetition of concepts during a speech, difficulty finding the right word;
7. Not finding things or losing the ability to retrace one's steps: inability to retrace one's steps to find objects left in unusual places;
8. Reduced or poor judgment: changes in judgment and decision making that manifest themselves in poor money handling skills and poor personal hygiene and care.
9. Withdrawal from work or social activities: giving up hobbies, social activities, work projects or sports activities;
10. Changes in mood and personality: appearance of psychological traits such as confusion, suspicion, depression, fright, anxiety[9].

However, it is important to note that the disease may begin long before clinical signs become apparent, making it critical to focus on developing techniques for early detection. Indeed, one of the main goals of research is to identify early signs that can distinguish healthy

aging from dementia. In this context, the concept of Mild Cognitive Impairment (MCI) was introduced, which represents an intermediate stage between normal aging and full-blown dementia, characterized by mild cognitive difficulties, but not enough to severely impair daily functions.

1.2.3 Diagnosis of AD

The diagnosis of Alzheimer's disease (AD) is based on the criteria of the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV-TR), and the guidelines of the National Institute of Neurological Disorders and Stroke-Alzheimer Disease and Related Disorders (NINCDS-ADRDA) working group[12]. The diagnostic process has two stages: initially a dementia syndrome is identified, followed by the application of AD-specific criteria. According to the DSM-IV-TR, the diagnosis requires the presence of a memory disorder and a deficit in at least one other cognitive domain, resulting in interference with activities of daily living (ADLs). Impairment in ADLs is considered the threshold parameter for the diagnosis of dementia, in addition to the mere presence of cognitive abnormalities. The NINCDS-ADRDA criteria, on the other hand, do not require impairment in social or occupational functioning, but specify that the onset of AD must be insidious and that there must be no other systemic or brain pathology that can explain the progressive cognitive impairment. The current criteria allow a probabilistic diagnosis of AD in clinical settings where definitive diagnostic biomarkers are not available. A definitive diagnosis of AD, however, can only be made with histopathological confirmation. Since 1984, when the NINCDS-ADRDA criteria were published, considerable progress has been made in the biological understanding of AD, allowing a more detailed view of the disease process. Today, the clinical phenotype of AD can be defined more precisely, thanks to the identification of distinctive markers of the disease. These include structural changes observable by magnetic resonance imaging (MRI), which show early involvement of the medial temporal lobe (MTL), changes in molecular neuroimaging by PET, which show hypometabolism or hypoperfusion in temporoparietal areas, and changes in cerebrospinal fluid biomarkers. A key aspect of this progress is the increasing focus on characterizing the early stages of AD, which precede dementia proper, defined by functional loss. It is crucial to distinguish between prodromal AD and normal aging, the latter considered as cognitive decline beyond simple aging. The term "mild cognitive impairment" (MCI) is most commonly used to refer to the disorder in

individuals who present with subjective and objective symptoms of impaired memory or other cognitive function, but without significantly impairing daily activities. Progression from MCI to dementia is more frequent than in the healthy population, but it is not an inevitable outcome. A more precise definition of AD is still needed for reliable diagnosis in the earliest stages of the disease. In summary, the NINCDS-ADRDA criteria are more focused and specific to Alzheimer's disease, while the DSM-IV-TR provides a more general definition of dementia that can be applied to different diseases[12].

1.2.3.1 Neuropsychological profile of AD

Anterograde memory represents the first aspect of memory to be impaired in Alzheimer's disease (AD), due to volume reduction in medial temporal structures and disturbances in connections between the hippocampus, amygdala, and entorhinal cortex. To distinguish between the mild stage of Alzheimer's disease (with MMSE score $\geq$ 24) and healthy subjects, one of the most relevant aspects is the tendency to remember mainly the most recent information, while long-term memories tend to be preserved. Tests such as those on verbal fluency reveal the difficulties in patients with mild Alzheimer's, as the temporal cortex is one of the first areas to be damaged. Executive functions, which include planning, organizing, and regulating behavior, are also impaired but at a more advanced stage of the disease, when frontal lobe changes become evident. Because these functions require interaction between different brain areas, deficits can occur even before there is overt damage in frontal regions. Indeed, patients with Alzheimer's disease show difficulty in directing attention between various tasks and in performing tasks that require the simultaneous use of multiple skills, as in the case of dual-task tasks. A typical example is difficulty remembering who said what during a conversation. Early diagnosis of Alzheimer's disease can be facilitated by taking into account language difficulties. Recent studies have observed that pathological individuals have problems finding words, a phenomenon that could be related to semantic memory deficits, making patients vague and imprecise in speech. Speech comprehension and the content of conversations are other areas that undergo changes as the disease progresses. The National Adult Reading Test (NART) is an instrument used to study how dementia affects reading skills. It has been observed that while reading comprehension and oral speech decline, the test is not particularly useful in the early stages of the disease. Diagnostic test batteries try to balance cost, time, and the number of cognitive functions to

be assessed. The choice of instruments also varies according to the stage of the disease being diagnosed. In mild Alzheimer's research, the priority is a sensitive diagnosis of anterograde memory and semantic memory. In the case of Mild Cognitive Impairment (MCI), the Mini-Mental State Examination (MMSE) is the most commonly used test to distinguish between individuals with dementia and healthy individuals. In addition, it allows classification of the severity of the condition: scores greater than or equal to 24 indicate a "very mild" condition, while scores below 10 suggest a "severe" form, those between 11 and 21 are indicative of a "moderate" condition, and above 21 correspond to a "mild" level, as reported by Haxby et al. [13]. To assess neuropsychiatric symptoms (NPS), the Neuropsychiatric Inventory (NPI) is often used, which includes 12 areas, each with a score based on frequency (0-4) and severity (0-3). The overall score is calculated by multiplying these two values and summing the scores obtained in the various areas. In addition to the MMSE, Rey's Auditory Verbal Learning Test, Digit Span (forward), Stroop Test and Trail-Making Test are also administered to measure attention. The Frontal Assessment Battery and the Digit Span (backward) are used to assess executive functions, while the Space Perception Battery and the Visual Object test explore visuospatial abilities[14]. Another useful diagnostic tool is the ADAS-Cog (Alzheimer's Disease Assessment Scale-Cognitive section), which takes about 30-40 minutes and assesses the cognitive domains most affected by AD, providing an overall error score. Regarding the differential diagnosis between Alzheimer's and vascular dementia, the Hachinski Ischemic Scale (HIS) can be a valuable support. It consists of 13 items, which are assigned 1 or 2 points depending on the condition (e.g., depression, hypertension, stroke). A total score of less than 4 points is indicative of dementia. Finally, according to a study by Storey et al., 2001[15], it is estimated that an average of about 2.6 ± 1.4 years elapses from the time the first symptoms appear to the first assessment.

1.2.3.2 Medical Imaging Exams

Medical imaging examinations, with an emphasis on functional neuroimaging, are used to diagnose the condition of aMCI (Mild Cognitive Impairment) or Alzheimer's disease. The first step is to rule out possible secondary causes of decreased cognitive performance, such as vitamin B12 deficiency, folic acid, or thyroid hormone dysfunction, as well as laboratory tests such as CBC or kidney function[4]. For second-level analysis, imaging techniques such as positron emission tomography (PET) and single-photon emission computed tomography

(SPECT), which employs technetium-99, are used. These tools make it possible to differentiate the disease in the early stages of diagnosis, as well as to frame neural damage at the onset of symptoms. PET is particularly useful in detecting the deposition of amyloid-(A) peptides in plaques in the living brain. However, the use of amyloid PET in clinical practice is still limited by its high cost, which is not covered by most health insurers. A more invasive but less expensive alternative is to examine cerebrospinal fluid (CSF) for A42, hyperphosphorylated tau peptide (p-tau), and total tau protein content. This method, although less expensive, has slightly lower diagnostic accuracy (85-90%) and carries the risks associated with lumbar puncture, with extended waiting times for results due to the scarcity of specialized laboratory facilities. However, a direct comparison showed that there is no significant difference in diagnostic accuracy between the A42:p-tau ratio in CSF and amyloid PET biomarkers[13]. Regarding the assessment of connections between brain areas, resting-state functional magnetic resonance imaging (fMRI) allows the observation of low-frequency trends (below 0.08 Hz). In a study by Wang et al. (2006)[16], it was shown that Alzheimer's patients exhibit a loss of neural synchronization, particularly between the right prefrontal areas and the right hippocampus. Comparing the connectivity of the right hippocampus between Alzheimer's patients and healthy individuals using 3D T1-weighted images, a significant reduction in connectivity is observed in the pathological conditions. In particular, regions such as the dorsal medial prefrontal cortex (MPFC), right wedge, and left wedge show a substantial decrease in connectivity in patients compared to healthy individuals. In addition to these techniques, magnetic resonance imaging (MRI) is a key tool for diagnosing and assessing the progression of Alzheimer's disease. MRI allows detailed imaging of brain structures and, in the case of Alzheimer's, identification of atrophy of crucial brain areas, such as the hippocampus and entorhinal cortex, which are among the first to be compromised by the disease. Atrophy in these areas can be monitored over time, providing crucial information to assess the progression of the disease. The combination of imaging techniques, such as PET, fMRI, and MRI, provides a comprehensive overview of brain changes associated with Alzheimer's, improving diagnostic accuracy and the ability to monitor changes over time.

1.3 Mild Cognitive Impairtment

Mild Cognitive Impairment (MCI) has become an important area of research as it provides an opportunity for early intervention in the degeneration process, potentially delaying the onset of neurocognitive decline, preserving the patient's quality of life, and reducing the substantial societal costs associated with neurodegenerative diseases.11 The modern definition of MCI emerged during the 2003 Key Symposium in Stockholm, where the official diagnostic criteria were refined and published in 2004. These criteria included the presence of cognitive complaints reported by the patient or caregiver, confirmation of objective cognitive impairment through neuropsychological testing, maintenance of independence in daily activities, and no presence of dementia. The National Institute on Aging and the Alzheimer's Association (NIA-AA) later made further refinements, but the core definition established during the Key Symposium remained largely intact. Deficits in MCI are usually mild, making it difficult to distinguish from other conditions, which leads to diagnostic inconsistencies. While memory is typically the first cognitive function to be affected, other areas such as language, visuospatial skills, executive functions, and reasoning can also be involved in some cases. According to the Mayo Clinic and Key Symposium definitions, MCI patients are categorized into two primary subtypes based on memory function:

- a-MCI: where only memory is impaired
- na-MCI: where memory is not involved, but language or visuospatial abilities are impaired.

Approximately 29% of individuals with MCI may develop neuropsychiatric symptoms such as depression, apathy, irritability, agitation, and abnormal motor behavior, with variations depending on the MCI subtype. Patients with a-MCI may experience apathy, agitation, and irritability, while those with na-MCI are more likely to exhibit depression and anxiety, both of which impact cognitive function[17]. Patients with MCI typically experience difficulties with anterograde episodic memory, meaning they may struggle to remember newly acquired information. Working memory and prospective memory, which are used for remembering future tasks, may also be impaired, leading to challenges in performing daily activities. These early cognitive impairments may suggest a future progression toward Alzheimer's dementia[18]. The prevalence of MCI among adults aged 71 and older is approximately

22%, and up to 40% of those who receive help with memory tasks exhibit signs of MCI. However, not all individuals with MCI progress to dementia.17 Research has shown that the conversion rate to dementia is around 18%, and MCI patients are 3 to 5 times more likely to develop dementia compared to the general population[19]. Further studies suggest that patients with multiple domain MCI are at a higher risk of developing dementia than those with a-MCI. For example, a study conducted at the California Alzheimer's Disease Center tracked a cohort of MCI patients over three years, finding that 65% of them progressed to some form of dementia. Among them, those with na-MCI were more likely to develop vascular or frontotemporal dementia, while those with a-MCI had a higher likelihood of progressing to Alzheimer's disease[20]. Finally, various etiological factors contribute to each MCI subtype, and by correlating these factors with expected degenerative pathologies, it is possible to predict the type of dementia a patient may develop in the future.

1.3.1 Diagnosis of aMCI

Research on aging and dementia is progressing rapidly toward more effective methods for early identification of clinical symptoms. The primary goal of clinicians is to detect individuals who are still asymptomatic but at risk of developing dementia. In this context, mild cognitive impairment (MCI) has been recognized as an intermediate stage between normal age-related cognitive decline and the earliest signs of Alzheimer's[21]. Early diagnosis of MCI is crucial, as this predescent stage has been shown to respond to pharmacological treatments that can, if not stop, at least slow progression to full-blown dementia.

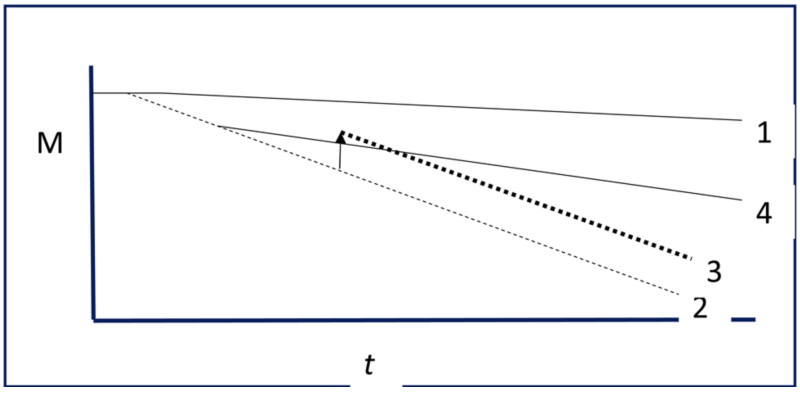


Figure 1.3: Depiction of the rate of memory decline over time in different conditions.

The image 1.3 shows the rate of memory decline (M) over time. In normal aging, memory decline occurs slowly (1). In contrast, Alzheimer's disease is characterized by more rapid cognitive decline, often beginning at a younger age (2). Current therapies improve cognition without altering the rate of disease decline (3). The expected effect of innovative therapies is a reduction in the rate of cognitive decline (4)[22]. However, diagnosis at an early stage is complex: impairment of a single cognitive domain often overlaps with what might be considered normal aging, with no obvious signs of functional deficits or behavioral changes. A diagnosis of MCI attributable to AD requires evidence of a decline, as indicated by:

- A change in cognitive abilities compared to previous levels, reported by the patient, a family member/caregiver, or identified through clinical judgment.
- Cognitive impairment in at least one domain, not necessarily limited to episodic memory, relative to age- and education-matched normative values. Impairment may also be observed in multiple cognitive domains.
- Preservation of functional independence, although mild problems in performing instrumental activities of daily living (IADLs) are allowed, even if the individual requires some level of assistance.
- No dementia, as this would imply more significant impairment in functional abilities.
- A clinical presentation consistent with the phenotype of AD, with no other potentially dementing conditions responsible for the symptoms.

The degree of diagnostic confidence can be strengthened by the presence of biomarkers:

- Maximum reliability when both an A accumulation biomarker and a neurodegeneration biomarker are present.
- Moderate reliability when an A biomarker is positive but no evidence of neurodegeneration is found[23].

To date, the most studied and validated biomarkers are Abeta42 and tau in the cerebrospinal fluid (CSF), hypometabolism on fluorodeoxyglucose positron emission tomography (18F-FDG PET), hippocampal volume (HV) on magnetic resonance (MR), and brain amyloid

deposition on amyloid imaging with PET. Hampel H, Burger K, Teipel SJ et al (2008) Core candidate neurochemical and imaging biomarkers of Alzheimer's disease. NIA-AA (National Institute on Aging and the Alzheimer's Association) defines two ways to detect A deposition:

- Monitoring of the level of A42 in cerebrospinal fluid (CSF)
- PET imaging

For neuronal degeneration, instead, three types of strategy are listed:

- Measuring the concentration of tau protein levels in CSF
- Identifying brain atrophy in medial temporal lobe or hippocampal volume through MRI
- Detection of low glucose metabolism in PET imaging involving temporoparietal regions.

Last edits on criteria dates to 2013, when the new version (V) of DSM was published. It consists of 2 steps: first of all, it's necessary to distinguish between normal functions, neurocognitive disorders (NCD), mild NCD, major NCD (dementia). Later, they characterize the aetiology of the problem. The mild NCD condition is equivalent to the definition of MCI from the Mayo Clinic. This table summarizes the core clinical criteria used to define mild cognitive impairment (MCI) according to four major frameworks: the original Mayo Clinic criteria, the Expanded/Key Symposium criteria, the National Institute on Aging-Alzheimer's Association (NIA-AA) workgroup criteria, and the DSM-5. Thè figure 1.5 summarizes the core clinical criteria used to define mild cognitive impairment (MCI) according to four major frameworks: the original Mayo Clinic criteria, the Expanded/Key Symposium criteria, the National Institute on Aging-Alzheimer's Association (NIA-AA) workgroup criteria, and the DSM-5.

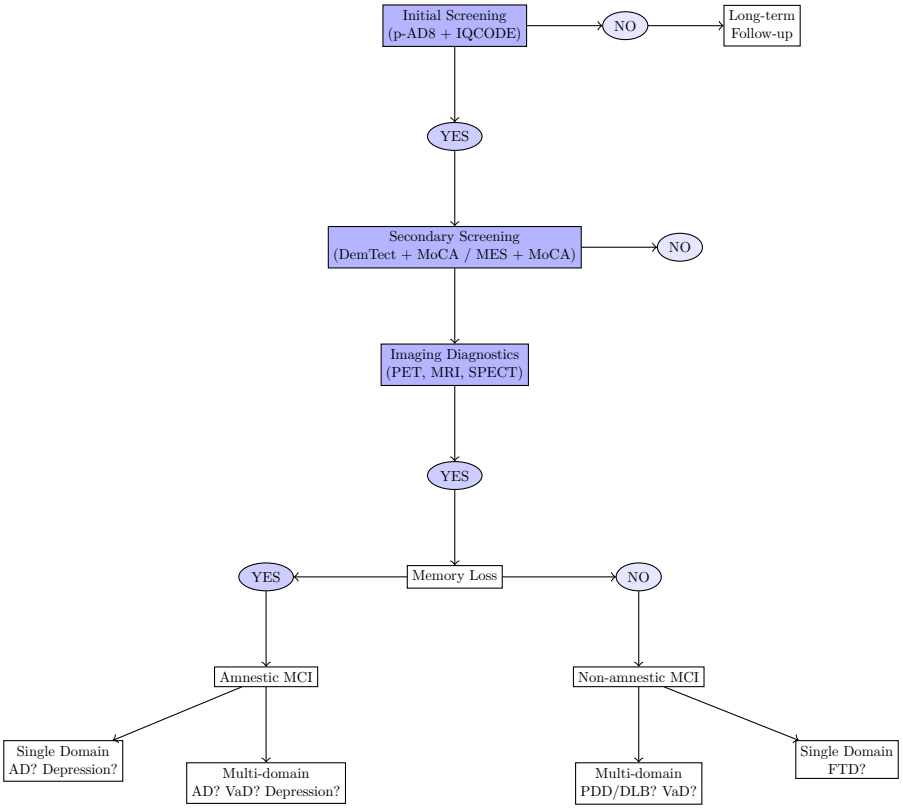


Figure 1.4: Flowchart for the Screening and Diagnosis of Cognitive Impairment[24].

CRITERIA	Ordinal Mayo Clinic	Key Symposium	NIA-AA	DSM-5
Self- or informant-reported memory complaint	X			
Self- or informant-reported cognitive complaint		X	X	X
Objective memory impairment	X			
Objective cognitive impairment		X	X	X
Essentially preserved general cognitive functioning	X			
Preserved independence in functional abilities	X	X	X	X
No Dementia	X	X	X	X

Figure 1.5: Clinic characterization of MCI. NIA-AA, National Institute on Aging-Alzheimers Association workgroup; DSM-V, fifth edition of the Diagnostic and Statistical Manual of Mental Disorders; X: criterion required[25].

1.3.1.1 Medical Imaging Exams

Recent advancements in medical imaging have significantly enhanced our ability to obtain precise anatomical, structural, and functional insights into the brain. These techniques complement neuropsychological evaluations by offering high-accuracy assessments of pathological changes, particularly in the context of Mild Cognitive Impairment (MCI) and its progression to Alzheimer's Disease (AD). PET imaging is widely used to detect amyloid-(A) deposition, a key biomarker in the neurodegenerative process. Using radioactive tracers such as Pittsburgh Compound-B (PiB) and Fluoro-2-deoxy-D-glucose (FDG), PET can precisely localize A deposits. Studies have shown that PET can achieve sensitivities as high as 96% and specificities reaching 100% in detecting these abnormalities[26]. Location of the protein can be identified through two different radioactive tracers:

- FDG-PET: In MCI patients, FDG-PET often reveals significant deficits in the temporoparietal regions and the posterior cingulate cortex—areas that are predictive of conversion to AD[27].
- PiB: This modality highlights increased amyloid deposition in regions including the posterior cingulate cortex, medial and lateral parietal lobes, frontal cortex, and lateral temporal lobe[25]. Due to its high cost, PET is primarily utilized in clinical research settings rather than routine clinical practice

While less expensive than imaging, CSF analysis is more invasive. It provides a window into the biochemical composition of the brain by measuring levels of tau protein and other biomarkers. Typically, elevated tau concentrations are observed in patients with MCI or AD compared to healthy controls. However, with turnaround times of several weeks and no single marker being universally accepted, CSF analysis remains a supplementary diagnostic tool (Schneider et al., 2009). Structural MRI is instrumental in detecting atrophy and other morphological changes, particularly within the medial temporal lobe—where hippocampal volume loss is a critical indicator of MCI and AD progression. Functional MRI (fMRI), which utilizes the Blood Oxygenation Level Dependent (BOLD) effect, offers superior temporal and spatial resolution compared to PET and does not expose patients to radiation. fMRI maps neural activity by tracking changes in blood oxygenation during specific cognitive or motor tasks. Notably, in AD, fMRI studies have documented reduced connectivity between

the right hippocampus and the posterior cerebral cortex, while structural MRI confirms volume loss in both the hippocampus and the posterior cingulate. cortex[25].

1.3.1.2 Neuropsychological Tests

Cognitive tests are employed to identify Mild Cognitive Impairment (MCI) and to evaluate the extent of cognitive deficits. It is generally observed that subjects with MCI score about one to one and a half standard deviations below the average expected performance. One of the most widely used assessments is the Mini Mental State Examination (MMSE), which takes roughly 10–15 minutes to administer. The MMSE comprises 11 steps that examine several functions—such as temporal and spatial orientation, immediate memory, attention, language, and basic arithmetic. The total score ranges from 0 to 30. According to Petersen (2004) in "Mild Cognitive Impairment as a Diagnostic Entity"[28], a score between 20 and 30 can indicate dementia; however, when daily functioning is largely intact, a score of 29 might still be considered normal. It is important to note that performance on this test can be influenced by the individual's level of education and literacy. In addition, the Milan Overall Dementia Assessment (MODA), which is structured in three parts, has been proposed as a more precise tool. For specifically detecting MCI, the Montreal Cognitive Assessment (MoCA) is often used. This test includes 30 questions designed to assess short-term memory, executive functions, attention, language, orientation, visuospatial skills, and concentration. Recent research is also comparing the Mini-Cog test with the MMSE. The Mini-Cog is much shorter—lasting about 3 minutes—and might be particularly suitable for MCI screening, although its outcomes, too, can be affected by educational factors. In the realm of behavioral analysis, several instruments are utilized. Among these are the Neuropsychiatric Inventory (NPI), the Cornell Scale, and the Geriatric Depression Scale (GDS). The latter two offer important insights into MCI. The GDS, for example, includes 30 items with categorical responses, yielding a total score between 0 and 30 (with higher scores indicating more severe depression), and a score around 3 is often associated with MCI or Alzheimer's Disease. Another comparable tool is the Clinical Dementia Rating (CDR, a value in a range from 0 to 3) scale, where a rating of 0.5 is indicative of MCI or very mild Alzheimer's Disease. This assessment, which requires input from a family member, evaluates memory, orientation, judgment, aspects of work and social life, hobbies, and self-care—with memory receiving the greatest emphasis—and classifies patients into

five levels, the highest of which corresponds to terminal dementia. For the amnestic subtype of MCI (a-MCI), diagnostic guidelines (as referenced in 2004mild) suggest that the first two steps of the diagnostic flowchart involve an evaluation of memory followed by an assessment of overall cognitive function, while the fourth step focuses on functional activities. Regarding functional evaluations, key measures include the Barthel Index, Basic Activities of Daily Living (BADL), Instrumental Activities of Daily Living (IADL), and Advanced Activities of Daily Living (AADL). The Barthel Index provides an estimate of a person's ability to perform basic tasks such as eating, walking, personal hygiene, and transitioning from sitting to lying down—with each task scored from 0 to 15 depending on the level of assistance required. Recent findings by del Carmen Díaz-Mardomingo et al. (2017) [29] have noted that individuals with MCI can exhibit deficits in IADL tasks, such as navigating public transportation. Originally developed for dementia evaluation, these tools are now also being used to differentiate between subtypes of MCI—specifically between amnestic MCI and multiple-domain MCI (which involves a broader range of deficits) versus single-domain non-amnestic MCI. The IADL test assesses abilities that most older adults typically use to maintain their independence, providing simple binary outcomes (dependent or independent). In contrast, the BADL is more reflective of physical abilities than cognitive ones. Finally, the AADL evaluates more complex and demanding skills, including hobbies, travel, and social participation. These advanced activities are influenced by mood and cultural factors as well; for example, reduced social engagement—particularly among women—has been linked to a decline in cognitive function. Neuropsychological assessments remain a cornerstone of MCI diagnosis due to their lower cost and ease of administration. However, these tests often suffer from issues of lower accuracy, high administration time, and variability in test protocols and cutoff scores, making cross-study comparisons challenging. In contrast, imaging modalities offer higher specificity and objectivity, albeit at greater expense and complexity. Additional neuroimaging techniques—such as SPECT, as well as volumetric MRI focusing on the hippocampus, and assays of biomarkers like A42 and tau in the CSF—have further contributed to our diagnostic capabilities. For instance, studies indicate that healthy individuals tend to have the largest hippocampal volumes, those with MCI have intermediate volumes, and patients with AD exhibit significant atrophy (Xu et al., 2000). Similarly, a reduced FDG uptake in the right temporoparietal cortex has been correlated with rapid progression from MCI to AD, often within an 18-month period.

1.4 Ocular Movement in AD

The focus of this thesis is the identification of eye movement patterns for the early diagnosis of Alzheimer's Disease (AD) through eye-tracker signal analysis. The main outputs of the eye-tracker signal include:

- **Fixations**: the period during which the eye remains relatively still on a point, measured in duration (ms) and spatial position;
- **Saccades**: rapid eye movements between two fixation points, characterized by amplitude, velocity, and direction;
- **Blinks**: frequency and duration of eye blinks, often used to analyze attention levels or visual fatigue;
- **Pupilometry**: variations in the pupil diameter, which can reflect cognitive and emotional states.

Regarding saccadic movements, patients with Alzheimer's Disease (AD) show several abnormalities, including fixation instability, prolonged latency for both reflexive and voluntary saccades, disturbances in the visual grasp reflex, and uncorrected errors in antisaccades. Another frequently observed phenomenon is hypometria, which correlates directly with the severity of the disease [30]. Numerous studies have highlighted an increase in latency and variability of saccades in patients with AD [31]. Furthermore, a higher number of uncorrected antisaccades and a lower number of correct antisaccades (directed toward the target), with a longer correction time, have been observed compared to healthy controls. These abnormalities seem to be linked to the impairment of the parietal and frontal cortices in AD, brain areas involved in saccade execution, antisaccade inhibition, visual attention, executive functions, and working memory, particularly in the dorsolateral prefrontal cortex (DLPFC) [32]. Performing antisaccades requires the ability to inhibit or correct movement toward the target, a task that necessitates maintaining the goal of the task in memory. Difficulties in executing antisaccades reflect the involvement of executive attention and the DLPFC, suggesting that the observed alterations in antisaccades are due to damage in these brain areas [33]. Regarding predictive saccades, no significant differences were

found between healthy subjects and those with AD, but greater variability among patients compared to controls was noted. This implies that patients with AD are able to perform anticipatory saccades, but they tend to be more dispersed around the target [32]. Regarding microsaccades, the study by Kapoula et al. (2014) highlighted that patients with AD experience alterations in the dynamics of these movements, such as duration, intersaccadic interval, and peak velocity, with a more oblique direction compared to healthy subjects, since microsaccades are normally horizontal relative to the fixation point. Moreover, regarding fixations, it has been observed that patients with Alzheimer's Disease (AD) often exhibit altered fixation periods. Specifically, patients tend to show unstable fixations with frequent micro-eye movements that disrupt concentration on the visual target. This fixation instability could be an indicator of the deterioration of the ability to maintain long-term visual attention, a process that mainly involves the parietal and frontal cortices, similar to what has been observed in saccadic movements and slow pursuit. The inability to maintain a stable fixation could reflect a deficit in attention and working memory, which are often impaired in AD patients [32]. Furthermore, abnormalities in fixations could significantly impact the quality of visual interaction with the surrounding environment, affecting the patient's ability to orient and respond to visual stimuli. Finally, a study by Ladas et al. (2014) observed that subjects with Mild Cognitive Impairment (MCI) showed a higher frequency of blinks compared to healthy controls [34]. This observation opens up the possibility of using blink analysis as an indicator of dopaminergic stimulation in MCI patients, although further studies will be necessary to confirm the effectiveness of this approach.

1.5 Aim of the Project

This thesis project follows the guidelines set by the OSR Ethics Committee of the Neurology Unit at IRCCS Ospedale San Raffaele. NAP4A2020 is an observational, low-risk longitudinal, non-drug spontaneous study protocol proposed by the Center for Advanced Technology in Health and Wellbeing in collaboration with the Neurology Unit at San Raffaele Hospital. The overall objective of the study is to observe how the NAP4A multisensory platform, a cognitive and behavioral assessment tool that includes wearable and non-wearable sensors, can support the clinician in early identification of the preclinical stage of AD by indicating the likelihood of risk of conversion to AD early. The analysis of aMCI conditions is carried out by integrating

the analysis of biomedical signals (electroencephalographic, photoplethysmograph, eye-tracker, facial expressions, and galvanic response) and neuropsychological assessments (performed by standard neuropsychological tests and the CANTAB test battery). The objectives of the study are:

- Primary objective: to develop a digital cognitive and behavioral assessment tool of rapid administration and interpretation for the study of the preclinical stage of Alzheimer's disease: NAP4A.
- Secondary objectives:
    1. Early identification of aMCI subjects at risk of conversion to AD through NAP4A;
    2. To create correlated indices of probability that the given aMCI subject may develop AD based only on data from the NAP4A platform.

The process involves basic steps:

- Data acquisition and dataset construction: using wearable and non-wearable sensors during neuropsychological tasks and visual stimulation, neurophysiological signals are collected and appropriately organized into a structured dataset using the Matlab platform.
- Feature extraction and selection: relevant parameters, derived from eyetraking, were identified to analyze cognitive and behavioral responses;
- Descriptive dataset analysis: data collected from signals were used to identify the eye patterns of each group (MCI and AD) and data collected from neuropsychological tests were used to perform longitudinal statistical analysis. Given the limited availability of recruited MCI patients, the initial focus is on the analysis of the control group.

Given the scientific literature, early diagnosis is crucial to slow the progress of the disease and reduce the social cost of AD.

1.6 Thesis Outline

The structure of this thesis is organized into six main chapters, each addressing a specific aspect of the research project. A brief overview of each chapter is provided below. The first chapter introduces the topic of neurodegenerative diseases, with a focus on Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI). This chapter provides a comprehensive overview of the clinical manifestations, diagnostic criteria, and current treatments for these conditions. The role of ocular movement analysis in understanding Alzheimer's Disease is also discussed. Chapter 2 describes the materials and tools used throughout the study. This includes an explanation of the study design, the participant sample, and the visual stimuli employed in the experimental procedure. The chapter also details the various tests used to assess cognitive function, including the CANTAB battery, as well as the technological platforms and sensors used for data collection, such as the eye-tracking devices and physiological measurement tools. In Chapter 3, the methodology adopted for this study is thoroughly described. This includes a detailed account of the experimental protocol, the preparations made for the CANTAB and NEFFIE phases, and the data collection process. Furthermore, this chapter explains the steps involved in biomedical signal recording and how the raw data was pre-processed and analyzed. Special attention is given to the signal filtering process and the extraction of relevant features for further analysis. Chapter 4 presents the preliminary results of the analysis. The findings from the CANTAB Battery Test and the eye-tracking signals are presented and interpreted. These results provide valuable insights into the cognitive and ocular responses of participants, offering an initial understanding of the potential correlations between eye movements and cognitive function in individuals with MCI and AD. - The fifth chapter discusses the results obtained in the study. The findings are compared with existing literature on Alzheimer's Disease and MCI, and the implications of these results are explored. This chapter also addresses the limitations of the study and suggests possible improvements in the methodology for future research. The discussion emphasizes the importance of ocular movement analysis as a potential diagnostic tool for early-stage dementia. The final chapter provides a summary of the key findings and conclusions of the research. It highlights the contributions of the study to the understanding of Alzheimer's Disease and MCI, particularly in relation to the use of eye-tracking technology. The chapter concludes with suggestions for future research, including the potential applications of ocular movement analysis in clinical settings and the

exploration of further diagnostic markers for neurodegenerative diseases.

Chapter 2

Materials

2.1 Study Desing

NAP4A2020 is a low-risk longitudinal, non pharmacological, spontaneous study protocol proposed by the Center for Advanced Technology in Health and Wellbeing in collaboration with the Neurology Unit of San Raffaele Hospital, under the leadership of Prof. Massimo Filippi. This interventional study is a case-control study and it is structured in three phases distributed over time (T0, T1 and T2) spaced by a period of six months. In the first phase of the study (T0), the subjects in experimental groups C and P (i.e. aMCI group and healthy control group, respectively) will undergo a digitized cognitive-behavioral assessment. At this stage, each subject will have to perform:

- At first digital cognitive tasks (i.e. Cambridge Neuropsychological Test Automated Battery - CANTAB) during which neurophysiological data are recorded. The test battery is executed via a tablet, and at the same time as it is executed, biological signals are recorded using appropriate sensors connected to the platform. Each battery test corresponds to a specific session in the NAP4A platform interface, which must be prepared and started manually by the back-end researcher;
- In a second step, the subject must view images administered through the NAP4A platform during which neurophysiological parameters are recorded. This task can be divided into three subphases:
  - Training: the subject is subjected to visual stimuli (described in Section 2.3) and is prompted to express the degree of perceived pleasantness in observing the images. This phase, which is of short duration, is an initial attempt to get the actual mechanism of the test understood during the second encoding phase. In this phase, there are 4 images.
  - Enoding: the exercise from the training phase is repeated but with a longer duration because of the large number of stimuli to which the subject is subjected. In this phase, the number of images is 16.
  - Recall: the subject is asked to express by binary response (yes/no) whether the stimulus to which they are being subjected has already been observed in the previous encoding phase. In this phase, the number of images administered is 28.

The raw neurophysiological data, recorded by the sensors during objective cognitive tasks, make it possible to identify behavioral signs (i.e., arousal, relaxation, hypoactivation, changes in eye and zygomatic activity that can be correlated with AD patterns) and cognitive signs (i.e., reaction time, vigilance, emotion recognition, inhibition skills, visual memory) specific to aMCI subjects at risk of conversion to AD. From the collection of these data, sensor-extractable parameters (i.e., probability indices) can be identified through AI analyses that will be able to predict subjects at risk of conversion to AD. In order to verify the reliability of the correlated probability indices of NAP4A in predicting cognitive impairment and behavioral change that can be associated with AD, the trial on the NAP4A platform is repeated at 6-month (T1 phase) and 12-month (T2 phase) intervals on both recruited groups. Only after obtaining a complete study and characterization of the parameters and reactions of each group will it be possible to compare them and validate the NAP4A platform as a digital tool for cognitive and behavioral assessment of the preclinical stage of Alzheimer's disease. The execution of the protocol is followed by a heterogeneous team of researchers consisting of neuropsychologists, engineers and computer technicians who guide, support and monitor each experimental session carried out at the office of the Center for Advanced Technologies in Health and Wellness.

2.2 Experimental Sample

The experimental sample to be recruited consists of 50 participants, aged 55 to 80 years. About 40% of the subjects will have amnestic mild cognitive impairment (single- or multiple-domain aMCI), while the remaining 60% of the subjects will serve as a healthy control group (HC). Patients are recruited from those visiting outpatient clinics specializing in dementia management or from patients admitted to the Neurology Unit, IRCCS San Raffaele Hospital. The control group consists of volunteers dedicated to contributing to the progress of research. The recruitment process of both groups follows specific and predefined guidelines. Inclusion criteria for single - or multiple-domain aMCI patients are as follows:

- Age between 55 and 80 years;
- Acknowledgement and signing of informed consent;

- Cognitive deficits reported by family members, physicians or the subject himself/herself;
- Absence of functional impairment or mild impairment of instrumental activities (to verify satisfaction of this inclusion criterion, the caregiver may also be asked to complete a questionnaire);
- Cognitive efficiency in the normal range (MMSE [1] score within 0.5 standard deviations of the mean value for subjects of the same age and educational level);
- Specific memory-related cognitive deficits or poly-sectional cognitive deficits (affecting memory and other cognitive areas) abnormal for age (1.5 standard deviations below the mean for control subjects of the same age and schooling) found by objective, standardized cognitive tests.

Exclusion criteria for aMCI patients are:

- Presence of major sensory deficits (hearing loss or hypovisus);
- Current or previous psychiatric illness;
- Presence of diagnosis of dementia or other neurological disorders;
- Presence of systemic diseases of history;
- Presence of cerebrovascular events in history;
- Use of alcohol or psychotropic substances in history.

These exclusion criteria are also valid for the group of control subjects. The inclusion criteria for the latter, however, are:

- Age between 55 and 80 years;
- Acknowledgement and signing of informed consent;

[1] MMSE: Mini-Mental State Examination: Neuropsychological testing for the assessment of intellectual efficiency disorders and the presence of cognitive impairment

- Absence of functional impact or mild impairment of instrumental activities;
- Global cognitive efficiency in the normal range (Mini Mental State Examination scores equal to or greater than 28);
- Absence of specific cognitive deficits (affecting memory and other cognitive areas) abnormal for age, sex, and schooling (1.5 standard deviations below the mean of control subjects of equal age and schooling) found by objective, standardized cognitive tests.

The selection of participants according to the above inclusion and exclusion criteria takes place during the screening phase, which is better defined in Section 2.4. The total number of subjects currently enrolled is 38, divided into 30 control subjects and 8 aMCI patients, respectively. The longitudinal study was completed by 22 control subjects and 3 aMCI patients.

2.3 Visual Stimuli

At the basal phase, each participant brings personal images representing significant moments in his or her life. During the testing phase, these images are alternated with NEFFIE images to analyze neurophysiological reactions to different stimuli and compare them between control and aMCI subjects. Variables are studied by asking the person to indicate the level of pleasantness of the image, or by assessing mnemonic abilities by subjecting the same images in different order and at different stages. During subsequent stages of the test, personal images are alternated with NEFFIE images, but each participant's own set of images remains the same. This allows for consistent assessment of neurophysiological reactions and mnemonic abilities over time, without introducing variations in personal content. NEFFIE images, extracted from engineer Alberto Sanna's book, are designed to stimulate the subject's cognitive and emotional functions by engaging them in complex situations and multiple realities. The term stands for Neuro - Photo - Aesthetic and refers to a type of photography characterized by an elaborate and dynamic structure that requires a high level of attention to interpret. These images present multiple realities within the same composition, inviting the viewer to explore different perspectives. Also referred to as "metapictures," these photographs capture unique moments in which a fleeting image is

reflected on a transparent or mirrored surface, giving rise to unexpected visual combinations and unexpected semantic connections. This effect stimulates both the eye and the mind of the viewer, making the perceptual experience richer and more engaging.

Figure 2.1: Example of Neffie image from Alberto Sanna's library. Neuro-Aesthetic photography is a visual language that would like to involve observer and its feelings

2.4 Screening Test

As indicated by the NAP4A2020 protocol, this project includes a screening phase, conducted by a neuropsychologist, to assess the subject's eligibility to enter the experiment, following the inclusion and exclusion criteria established by the protocol through the participant's medical history and the administration of the neurological rating scales:

- Assessment scales of daily living skills: activities of daily living (ADL, Activities of Daily Living) and Instrumental Activities of Daily Living (IADL). These scales measure functional ability. The ADL scale allows assessment of functional changes

in elderly patients such as cognitive and physical decline associated with aging or neurological disorders such as dementia.[35] The ADL scale is an instrument used to assess a person's independence in performing basic daily activities. These activities include essential personal care tasks such as: washing, dressing, using the bathroom, and feeding oneself, and functional decline can be shown by the inability to perform these tasks alone. The scale assigns scores ranging from 0 (dependent) to 6 (fully independent) based on the patient's ability to perform each task.[36]

- The IADL scale is designed to assess the ability to perform complex activities in the elderly or in people with chronic or neurological conditions. It is a clinical tool used to assess a person's autonomy in performing daily instrumental activities considered necessary for maintaining independence. It consists of a list of eight complex functions that require competence in the use of tools. There are 8 activities considered: household management, meal preparation, money management, use of telephone and other devices, taking medications, use of transportation, laundry, and ability to shop. [37] The scale provides a score for each activity and, generally, the values are between 0 (dependent) and 8 (completely independent). These tests, in the case of aMCI subjects, are administered to the caregiver. [36]

- Screening scale for cerebrovascular events: the Hachinski Ischemic Score (HIS) is a clinical tool used to discriminate vascular dementia and primary degenerative dementia (such as Alzheimer's disease). It was developed by Hachinski Vladimir and consists of a set of criteria that assigns scores based on the patient's clinical characteristics. Each criterion is assigned a score (positive or negative), and the total score varies in the range of 0 to 12. A score less than or equal to 4 suggests primary degenerative dementia. A score greater than or equal to 7 indicates probable vascular dementia. An intermediate score of 5 or 6 indicates an area requiring further investigation.[38]

- Global cognitive efficiency screening test: the Mini Mental State Examination (MMSE) is an instrument used to assess cognitive function. It consists of 30 questions that test skills such as: orientation and knowledge of time and place, memory (word repetition and recall), attention and computation, language (naming, repetition and comprehension), and visuoconstructive skills through drawing. The MMSE score cannot be considered evidence of dementia per se but may signal the possibility of its presence.[39] The score ranges from 0 to 30 and reflects the subject's level of cognitive function:

  - Score 24-30: normal cognitive function (may vary with age and education level);
  - Score 18-23: mild cognitive decline (possible indication of dementia);
  - Score < 17: severe decline (indicative of advanced dementia).[40]

  In 20% of cases, people with dementia achieve a higher score, indicating the absence of the condition. This may be due to the level of schooling or the stage of dementia. It has been found that the MMSE is not sensitive to the aMCI as it assesses global cognitive efficiency.[40]

- Battery of standardized cognitive tests assessing different cognitive areas: memory, executive functions, visuospatial skills, language, processing speed, social cognition. For each cognitive domain, different neuropsychological tests established by the NAP4A2020 protocol are performed.NAP4A2020.

1. Memory

   (a) Rey Auditory Verbal Learning Test: a neuropsychological test designed to assess verbal memory and short-term verbal learning. The test consists of a list of 15 words for the patient to listen to and repeat in several stages. The test investigates various aspects of memory such as immediate memory, retentive memory and long-term memory.[41]

   (b) Three objects and three places test: a neuropsychological test used to assess spatial memory and long-term memory. The test consists of asking the patient to remember and repeat three objects and three places that are shown at the beginning of the test. [36]

   (c) WAIS - IV Digit Span Forward: a test that assesses short-term memory in that the subject is asked to repeat the sequence of numbers previously described by the examiners. As the answer is correct, the sequence is increased by one and the span will be the longest number of digits the subject can remember. [36]

   (d) Rey's complex figure: a test to assess visual memory during which the subject is asked to redraw a figure he or she is observing. Then, the subject has to reproduce the figure without the visual stimulus, and a score is associated with the realized image. [36]

(e) Span Memory: an exercise to assess the amount of visuospatial information gathered in recent memory. The examiner asks the subject to touch a series of cubes following the order he or she has just shown.[36]

2. Executive functions

   (a) Trial-making test: requires you to connect 25 elements that can be just numbers, or letters and numbers following an ascending order. [42]

   (b) Stroop test: a test that aims to observe selective attention during which the neuropsychologist asks the subject to say the color of the word they see, considering that words are names of colors.[43]

   (c) Latter fluency test: studies the activities of the frontal region and is an indicator of neurodegenerative decline in the disease and belongs to the category of verbal fluency tests. Subjects have to generate a series of words, as long as possible, starting with the same letter.[44]

   (d) Digit Span Forward: the repetition of a sequence of numbers in reverse order is required. Usually the result varies between 5 and 8 points. As the participant repeats the sequence correctly, the number of digits is continued to increase.[45]

3. Visual-spatial skills

   (a) Copy of simple drawings: the participant is asked to reproduce a simple drawing.[46]

   (b) Complex figure: the reproduction of a complex picture involving 18 graphic elements is asked.[46]

4. Language

   (a) CaGi: Italian battery for the assessment of semantic memory disorders and the detection of linguistic-semantic disorders

5. Processing speed Some WAIS-IV tests, measure processing speed by identifying figures and symbols:

   (a) The "Cifrario": asks people to associate symbols and numbers by investigating accuracy and speed.

(b) Symbol Search: asks people to detect whether a specific symbol is present among a set of symbols.[47]

6. Social Cognition

   (a) Story Based Empathy Task (SET): Subjects must choose the correct ending of a series of pictures, demonstrating the ability to understand intentions and emotions.[48]

   (b) Ekman 60 Faces Test is an instrument to assess the ability to recognize facial emotions. The pictures represent six emotions, and the maximum achievable score is 60 points (10 for each emotion).[49]

In addition to what has been mentioned so far, behavioral aspects are also investigated. Useful tools, provided by the protocol, in this area are:

- Geriatric Depression Scale (GDS) administered to subjects over 65 years of age. Consisting of 15 items to assess geriatric depression, it is not suitable for participants scoring less than 15 on the MMSE. Each question can collect 0 or 1 point. If the final score varies between 6 and 9, the possibility of depression exists, while for higher levels the problem is categorized as probable.[50]

- Beck Depression Inventory II (BDI-II) is appropriate for subjects under 65 years of age. Following DSM-IV guidelines, this test measures the severity of depression symptoms, delving into changes involving sleep, appetite, but also cognitive functions such as self-esteem. It requires the subject to answer 21 multiple questions. A partial score can be collected for each area and a total score for the entire test. Distress can be assessed if the score is above 91 points, and its severity increases with the total score.[51]

- State-Trait Anxiety Inventory (STAI) consists of 40 questions that take 10 minutes and essentially assess anxiety, differentiating it from depression, and a 4-point scale is applied to each item. It can be divided into 2 different scales (State Anxiety or temporary condition and Trait Anxiety or stable condition). A threshold of 39-40 points differentiates anxiety symptoms, although in older adults this is shifted to 54-55.[52]

- NeuroPsychiatric Inventory (NPI) measures the development of psychiatric symptoms throughout the duration of treatment. It is a test to investigate 10 areas of behavior (hallucinations, dysphoria, anxiety, irritability) plus two domains. Questions involve multiple or binary responses. The caregiver is asked to rate the frequency of occurrence of a given behavior and its severity. The total score is calculated by multiplying the frequency by the associated severity level. [53]

2.5 Cantab Battery Test

The Cambridge Neuropsychological Test Automated Battery (CANTAB), originally developed at the University of Cambridge and now available in commercial form through Cambridge Cognition, is a computerized cognitive assessment system consisting of a battery of neuropsychological tests administered to subjects using a touchscreen device[54]. The CANTAB test battery includes 25 tests that examine various areas of cognitive function including:

- General memory and learning
- Working memory and executive functions
- Visual Memory
- Attention and reaction time (RT)
- Semantic and verbal memory
- Decision-making capacity and response control

The CANTAB combines the accuracy and rigor of computerized psychological tests while maintaining the wide range of measures required by a neuropsychological battery. It is suitable for both young and old subjects, and aims to be independent of educational level and language through the use of nonverbal cues in most tests.[55] CANTAB can be administered by a trained assistant without the need for verbal instructions. Responses are given via the touch screen or a button, depending on the needs of the task. Each task begins with practice items at a basic level. The design and interface of the CANTAB tests aim to

minimize the effects of computer experience and computer-related anxiety. The software automatically measures and saves the parameters related to each task in a .csv format file in which the first columns provide information about the test subject, experimental stage, age, gender, and schooling level, and the subsequent columns refer to the parameters extracted from the individual tests.[56] To interpret and understand the numbers obtained, the range within which the parameters can vary is specified in a .pdf file, with descriptions, names, and variations of the parameters themselves (refer to the guide available at the following link:[57]). This phase of the experiment, referred to as the Cantab phase, consists of performing 6 consecutive tasks with a total duration of about 18-20 minutes. Throughout the execution of the test battery, biomedical signals and information of interest are recorded via EEG helmet, Shimmer and Webcam. The eye-tracker is not required at this stage but only during the execution of the previously described NEFFIE tests. The team involved in the implementation of the protocol selected the most suitable tests for the purpose of the research. This battery of tests can be customized, including or excluding some of the available tests that are deemed more or less relevant to the specific goal and scope of the research. Different aspects of cognitive function are assessed through distinct groups of tests, which can be divided into domains of cognitive interest. The "Attention and Psychomotor speed" domain takes into account the relationship between cognitive functions and physical movements and allows for the assessment of reaction time, movement time, and alertness. This makes it possible to identify an individual's ability to detect and respond quickly to changes in the external environment, such as the presence of a stimulus. The "Match to Sample Visual Search (MTS)" task is the one selected from this domain. It lasts about 7 minutes and assesses attention and visual search ability while balancing speed and accuracy. The participant is shown a complex visual pattern in the center of the screen, and after a short interval, other similar patterns appear around the edge of the screen. The participant must indicate, by selecting it, which of the patterns in the panes corresponds to the central one. The output parameters are 136 and provide a measure of cognitive abilities related to visual attention and short-term memory. The main ones include:

- Response time (RT): time taken by the participant to identify and select the correct pattern from the stimulus presentation;
- Accuracy: percentage of correct responses;
- Number of errors: count of the number of errors made during the test that could

indicate difficulty in maintaining attention or visual recognition.

- Visual search speed: speed at which the participant scans different visual stimuli.This is specific to the MTS test.

The first three parameters are made explicit in all tests to which the participant is subjected. The "Exectuve Function" domain assesses high-level thinking and decision making. It is divided into 4 subdomains: mental flexibility, planning, strategy and response inhibition. Multitasking Test (MTT) and Spatial Working Memory (SWM) were selected from this domain. The Multitasking Test has a duration of 8 minutes and assesses the participant's ability to handle conflicting information provided by the direction of an arrow and its position on the screen, while ignoring information not relevant to the task. The test features an arrow that can appear on either side of the screen (left or right) and can point in either direction (left or right). In each test, an indication at the top of the screen informs the participant which button to press, left or right, based on the "side the arrow is on" or the "direction it is pointing." In some parts of the test, this rule remains constant for all trials (single task), while in others it may vary from trial to trial randomly (multitasking). The flexible use of both rules requires more cognitive effort than the application of a single rule. There are 46 indicators derived from this domain that provide an overview of the cognitive skills involved in information management and multitasking. The main ones are:

- Response latency: time between stimulus presentation and participant's response;
- Effectiveness of multitasking management: ability to switch between rules and handle competing tasks, measured by how well the participant adapts to rule changes;
- Inhibition error: a parameter that measures the ability to ignore irrelevant information and suppress impulsive responses.

The Spatial Working Memory (SWM) lasts about 4-6 minutes and requires manipulation of visuo-spatial information. The goal of the test is to find among the various boxes shown on the screen a yellow token used to fill an empty column on the right side of the screen. Depending on the difficulty of the test, the number of boxes can increase to a maximum of 12. In addition, the color and location of the boxes change on each test to prevent participants from using habitual search strategies. There are 19 parameters derived and

provide information on cognitive abilities related to spatial working memory and strategic planning ability. The main ones include:

- Task difficulty: the SWM test progressively increases in difficulty as the participant progresses, adding more spatial targets to memorize and retrieve. Then difficulty can be measured according to the complexity of the trials completed without errors.
- Performance index: combines the number of errors, time taken, and correct answers to give an overall view of the participant's spatial memory ability.

The "Emotion and Social Recognition" domain assesses the ability to respond to emotionally charged stimuli. From this domain, the Emotion Recognition Task (ERT) test was selected that can assess the ability to recognize six basic emotions in facial expressions, considering different intensities of expression. The task takes about 6-10 minutes during which morphological images illustrating facial features of real people expressing a specific emotion are shown on the screen, one at a time. The image is displayed for 200 milliseconds and the participant must select which emotion is represented by the face from six options (happiness, surprise, disgust, anger, sadness and fear). The output parameters are 35 include the percentage of correct responses and the number of correct or incorrect responses, as well as the overall response latencies. Finally, the Memory domain includes sensitive, accurate and reliable digital cognitive tests to assess memory and monitor changes over time. The domain task used is Pattern Recognition Memory (PRM). The test has a duration of 4 minutes and is a visual pattern recognition test in a two-choice forced discrimination paradigm. The participant is presented with a series of visual patterns, one at a time, in the center of the screen. These patterns are designed so that they cannot be easily labeled verbally. In the recognition phase, the participant must choose between a pattern he or she has already seen and a new pattern. At this stage, the test patterns are shown in the reverse order of their original presentation. This process is then repeated with new patterns. The second recognition phase is administered after a waiting period, usually 10-20 minutes. There are 9 parameters extracted and include the number and percentage of correct trials, as well as latency (the speed of the participant's response). All information regarding the Cantab battery can be viewed on the official Cambridge Cognition CANTAB website.[45]

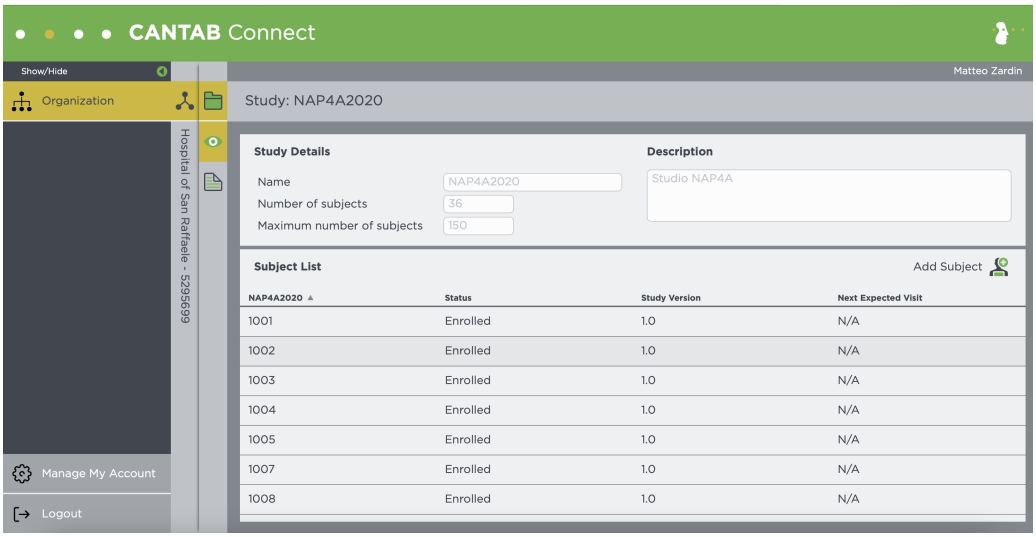


Figure 2.2: Data saving screen in the CANTAB Connect software for managing and archiving cognitive performance

2.6 NAP4A Platform

The NAP4A platform, employed by the NAP4A2020 protocol, is a digital multisensory platform consisting of a desktop software application (which allows simultaneous viewing of visual stimuli by the subject and recording of neurophysiological signals) and a set of wearable and nonwearable sensors for acquiring neurophysiological signals. During the experiment, following the protocol, the NAP4A platform is used in dual mode:

- During the first phase ("CANTAB phase") only the module intended for recording signals from the wearable and non-wearable sensors connected to the platform is used while the subject is asked to perform digital cognitive tasks (e.g. CANTAB);
- The second phase "NEFFIE" includes the full use of the NAP4A platform, i.e. the software application for administering a non-standardized cognitive task with visual input (e.g. photos) and for recording signals from the sensors. In this mode, personal and non-personal visual stimuli of various kinds (i.e. photos, postcards, paintings with biographical and non-autobiographical content and of recent and past events)

are administered to the subject through the software application while at the same time recording data acquired by wearable and non-wearable sensors connected to the platform.

The parameters of the acquisitions can be adjusted through the software interface that allows visualization of both the stimuli administered to the subjects and the signals acquired in real time. The latter can be visualized allowing monitoring of the quality of wearable sensor positioning and visualization of large artifacts due to subject movements.

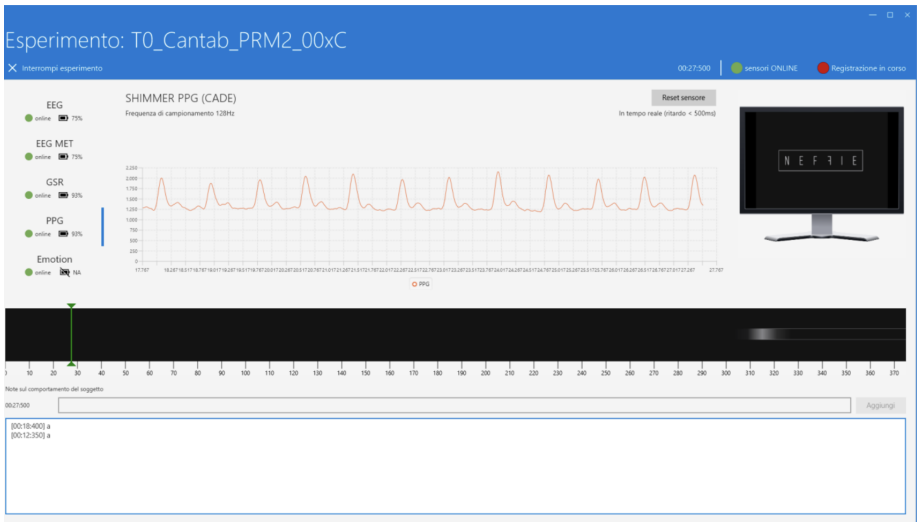


Figure 2.3: NAP4A Interface during acquisition

The interface makes it possible to visualize the activation and synchronization of the sensors during test execution (thanks to the presence of the online/offline indicators associated with different colors such as green or red). In case of technical issues related to the sensors, it is possible to reset the device via the "Sensor Reset" button directly on the interface. In addition, there is an alert indicator for possible issues such as sensor synchronization failure and low sensor battery level. From the interface it is possible to annotate, point to the "Subject Behavior Notes" section, the initial and final time instants of the various Cantab battery subtests by a back-end operator pressing a random key on the keyboard. It is important to note, at this stage, the timestamps related to the start and end of each task as they will be needed during the signal processing phase. Finally, information can be

derived about the quality of the EEG signal as a function of the percentage of electrode contact on the skin.

2.7 Sensors

During each session of the experiment, sensors useful for acquiring biomedical signals are adopted and used for subsequent parameter extraction. Each subject is equipped with SHIMMER GSR 3+ Unit (Shimmer Sensing Company), EEG Helmet INSIGHT-EPOC (EMOTIV Company), Eye-tracker EYE-T 4C (TOBII) and Webcam C922 Pro Stream 1080p.

2.7.1 Shimmer GSR3+ Unit

The Shimmer3 GSR+ unit is a wireless sensor that monitors skin conductivity and allows measurement of skin conductance (also known as electrodermal resistance or electrodermal activity) between two electrodes attached to two fingers of the same hand.In response to internal and external stimuli, sweat glands become more active, increasing the moisture content on the skin and allowing electrical current to flow more easily by changing the balance of positive and negative ions in the secreted fluid (increasing skin conductivity and thus decreasing skin resistance).The Shimmer3 GSR+ also provides an additional channel that can capture the optical pulse signal or PPG (plethysmograph) signal to estimate heart rate.It is designed to be wearable; in fact, the GSR+ unit is free of wired constraints and provides reliable data via a wireless transmitter.Reusable finger electrodes, disposable electrode pads and a wrist strap are included.Such a device for measuring skin conductance uses a known current that is injected into the applied part, so that by measuring via a voltmeter the voltage, precisely the desired magnitude is derived. A loading effect is present due to the internal resistances of the device. Actually, what the device provides is a voltage that is converted by an ADC to a 12-bit number, which represents just the skin conductance, according to the following equation:

$$R_s = \frac{R_f}{\left(\frac{\text{ADCv} \cdot 3V}{4095} \div 0.5V - 1\right)} \tag{2.1}$$

Where R_f is the value of the feedback resistance, in Ω, used by the sensor for a given interval, ADC_v is the 16-bit output of the ADC, R_s is the measured value of the skin resistance in Ω, $\frac{3V}{4095}$ is the voltage per bit of the ADC, where the reference voltage of the ADC is 3.0 volts and the 12-bit ADC has 4096 different measurement levels. Additionally, 0.5V is the reference voltage used by the sensor. Typical skin resistance ranges from 47 kΩ to 1 MΩ, and the GSR+ unit is designed to resolve skin resistance levels from 8 kΩ to 4.7 MΩ. It is recommended to use Ag/AgCl electrodes with a snap connector. The surface area of the electrodes should be kept to a minimum: 1 cm^2 is ideal. One electrode should be placed on the palmar surface of the medial phalanx and the other on the palmar surface of the distal phalanx of two neighboring fingers. The SHIMMER guide recommends setting a sampling rate greater than 100 Hz, and for the study, it is set equal to 128 Hz. [58]

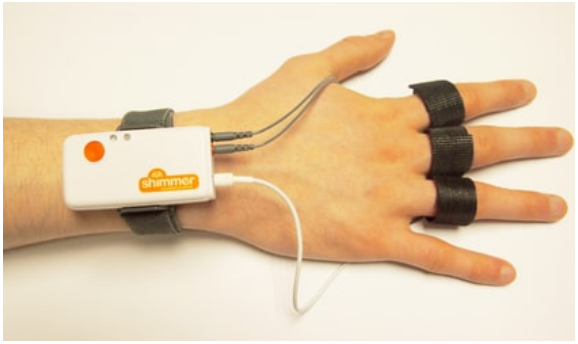
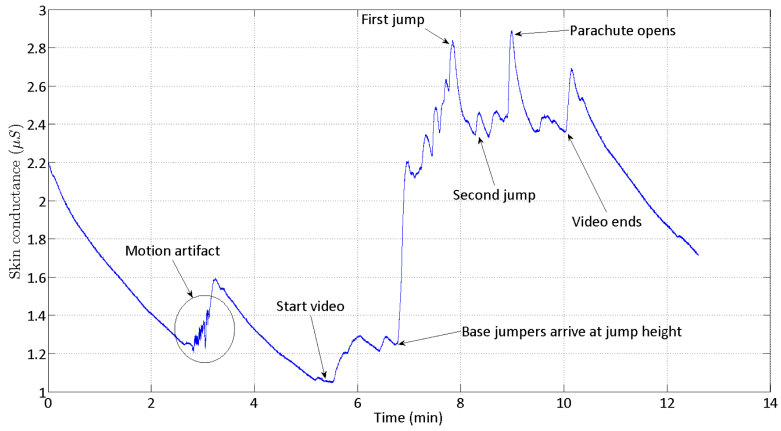


Figure 2.4: Example electrode positioning for GSR

2.7.1.1 Galvanic Skin Response Signal

The GSR signal is a skin resistance signal and has a characteristic waveform with a rise time of about 1-2 seconds and a longer fall time. The GSR signal is characterized by two components: a slower component called the tonic component and a component that varies rapidly over time called the phasic component. The two components originate peripherally but are mediated by the central nervous system and carry different information:

- Tonic component: expresses the absolute value of skin electrical resistance, is an index of the general state of activation of the body's nervous system. The tonic value is higher if the individual is calm and relaxed, but if the individual is agitated and



H

Figure 2.5: Example of GSR signal

nervous, skin sweating increases and skin electrical resistance is lowered. This tonic component is supposed to be related to wide-range brain activation, directly related to the action of norepinephrine and adrenaline produced by the adrenal medulla; in practice, it is related to sympathetic tone.

- Phasic component: represents the rapid responses provoked by purely emotional, sensory or ideational stimuli. It depends on the action of acetylcholine produced and disposed of locally. The tonic component is the low-frequency baseline of the signal (0-0.05 Hz) while the phasic component overlaps with the tonic and covers a frequency range from 0.05 to 1.5 Hz.[59]

As with any sensor that uses electrodes attached to the skin, where there is motion at the electrode site, there will be motion artifacts in the signals. This effect can take the form of a high-frequency noise component in the recorded signal if the electrodes are not firmly attached and lose contact with the skin during movement-even a partial loss of contact will cause this effect. Also, if the electrodes are pressed firmly against the skin, they will have better contact and, therefore, measure a higher conductivity value, compared to a looser connection. Usually, one tries to avoid these effects by securing the electrodes as securely as possible. It would also be advisable to avoid movements that cause the Velcro straps to friction, causing the electrodes to move. A low-pass filter should be applied to the data to remove high-frequency noise that can be attributed to motion artifacts and other noise components. A cut-off frequency of 1 to 5 Hz can be used without affecting the data of

interest due to the slowly variable nature of GSR responses.[58]

2.7.1.2 Photoplethysmogram Signal

Photoplethysmography is a noninvasive technique that allows continuous measurement of the blood flow and volume that occurs with each cardiac cycle. Blood volume decreases during the diastole phase and increases during the systole phase. A light-emitting mechanism that illuminates tissue and a detector that measures changes in light intensity as a function of changes in blood volume in the vessels are exploited to detect the PPG signal.

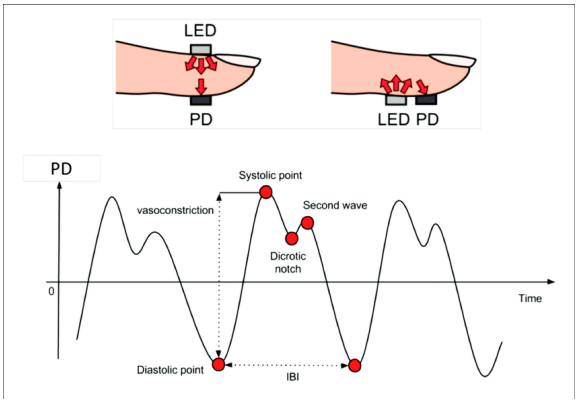


Figure 2.6: PPG sensor including the light source (LED) and the light receiver (PD)-two different setups can be used. The curve below illustrates a proto-PPG signal measured by PD.

The light emitted by the LED passes through biological tissue and can be absorbed by different substances such as skin pigments, bone, venous and arterial blood. An increase in the intensity of the captured light indicates a decrease in blood volume. The waveform of the PPG signal consists of a continuous component of the current (DC) and an alternating component (AC). The former is due to the presence of the tissues interposed between the sensor and the measurement point, the presence of venous blood and the nonpulsatile component of arterial blood. This component is influenced by respiration. The AC component depends on changes in blood volume during the systolic and diastolic phases of the cardiac cycle. To obtain a good PPG signal, it is necessary for the emitted light to have wavelengths in the visible or infrared because low wavelengths can be absorbed

completely by melatonin (wavelengths of 660 nm and 940 nm are optimal) because they have opposite absorption levels. The figure indicates the two modes of PPG operation: transmission or reflection. PD is denoted as the spectral response of the photodiode. In transmission mode, the PD, located opposite the LED, detects light transmitted through the tissue. In reflection mode, on the other hand, the PD detects light reflected from the tissue, bones or blood vessels. The waveform of the PPG signal has two phases: anacrotal phase and catacrotal phase. The first phase corresponds to systole, while the second phase corresponds to diastole.[60]

2.7.2 Emotiv™ Insight-Epoc Helmet

The EMOTIV Insight is a 5-channel EEG system with semi-dry polymer sensors that provide clean and reliable signals. Its sleek design and 20-hour battery life make it ideal for self-quantification, brain-computer interfaces, and field research. This sensor has been used to acquire the electroencephalographic signal used to study cognitive function and neuronal activity. Emotiv Insight interprets mood and emotions, such as the level of arousal, relaxation or stress, and basic mental commands, such as pushing, pulling, levitating, rotating. Insight's futuristic design is lightweight and less invasive to wear and requires no preparation as the sensors work dry, requiring no conductive gel or saline solution. The sampling rate is 128 Hz and includes 9 axis motion sensors working at 32 Hz. The sensor offers recordings on 5 channels, following the 10-20 positioning standard, in which the electrodes are positions in the locations: AF3, AF4, T7, T8, Pz to study the frontal, temporal and parietal lobes. This placement allows concentration of the study on the frontal lobe, which is responsible for reasoning, planning and intellect, and the temporal lobe associated with long-term memory, face recognition and object perception.[61]

Figure 2.7: Emotiv-Insight 5 Channel Mobile EEG

The EEG is a waveform that includes several frequency components, referred to as component bands or frequency bands because what distinguishes them is the frequency value. The different rhythms are represented in the table 2.1 with decreasing trend (starting from the fastest rhythms to the slowest rhythms).

Rhythm	**Frequency**	**Description**
GAMMA	30 - 50 Hz	Linked to moments of deep concentration, cognitive processing, and peak performance.
BETA	12 - 30 Hz	Associated with intense mental activity, physiological arousal, focus, and complex tasks.
ALPHA	8 - 12 Hz	Resting state, relaxed wakefulness, and meditative states.
THETA	4 - 8 Hz	Typically observed during drowsiness and the early stages of sleep.
DELTA	0.5 - 4 Hz	Characteristic of deep sleep, particularly in stages 3 and 4.

Table 2.1: Brain wave rhythms, their frequency ranges, and associated cognitive or physiological states.

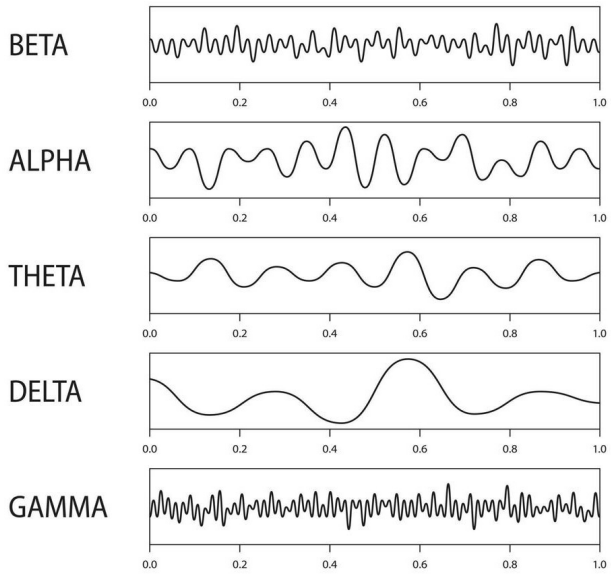


Figure 2.8: Graphical representation of brain wave patterns, illustrating the frequency ranges and characteristic activity of different brain wave types (Alpha, Beta, Theta, Delta, and Gamma).

2.7.3 Tobii™ Eye-Tracker

The eye-tracker signal is oculometric data that allows eye movements to be monitored and recorded with high accuracy. The eye-tracker is a device used to determine where a subject is focusing his or her gaze[62]. This paper considers the screen-based eye-tracker TOBII™ 4C, with a sampling rate of 90 Hz. Screen-based eye-tracking devices, such as Tobii's, generally consist of the following major hardware components (Figure 2.9):

- Infrared light illumination modules: to determine where the subject is looking, Tobii eye-trackers use a technique called Pupil-Center Corneal Reflection (PCCR) based on the use of light in the near-infrared (NIR) spectrum that generates reflections on the cornea and pupil.
- Camera sensors: eye-trackers are equipped with near-infrared light-sensitive camera sensors that capture images of the eyes and light reflections. The camera sensors are positioned in front of the user and have a clear view of the eyes. Depending on

the speed of the eye-tracker, information on specific types of eye movements can be captured.

- Processor (image detection, 3D eye model, gaze mapping algorithm): the eye image is captured by the camera and sent for analysis. Advanced image processing algorithms are used to estimate a 3D model of the person's eye and determine the location of the eye in space. During image analysis, the center of the pupil and corneal reflection are detected to calculate the fixation point. The fixation point, in the case of the Tobii eye-tracker, can be accurately calculated without the need to use a head support or chin rest[62].

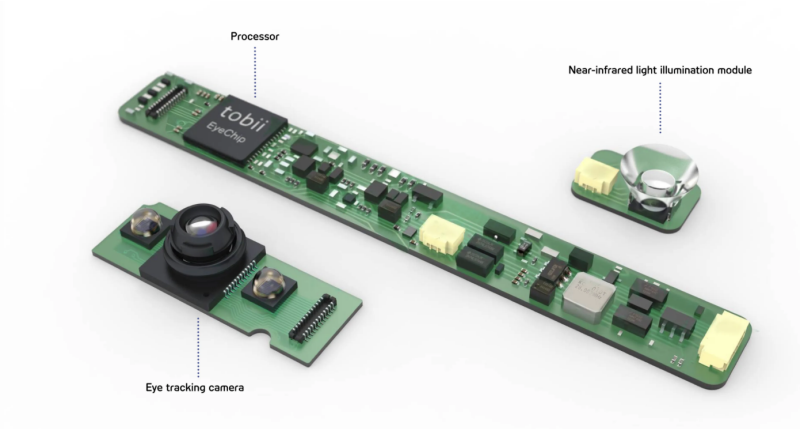


Figure 2.9: The key components of an eye tracker

The result consists of a pair of coordinates (x,y) identifying the fixation point and a timestamp corresponding to the time of acquisition. Eye movement can be described by fixations and saccades. Fixations refer to the pause period (0.2-0.6 s) over a region of interest, while saccades are rapid movements (20-100 ms) between fixations. The target can be either stationary or moving, and even in the case where a subject looks at an item moving in the region, it is considered to be a fixation. Intervals between fixations are considered saccades, and at those times the line of sight is shifted and directed to another point of interest. Fixations are derived by processing through a mathematical algorithm that associates a series of fixation points with an event, following a conversion based on specific features. By evaluating features such as the duration in time of fixation or the size

of the region explored, cognitive function, attention span and engagement of the visual stimulus can be studied[63].

2.7.3.1 Classification of eye-trackers

Techniques for acquiring gaze tracking can be invasive or remote. Invasive techniques involve direct physical contact between the device and the person using it (e.g., contact lenses or electrodes). Remote techniques, on the other hand, require the use of external devices such as cameras to obtain an image of the eye. Among remote gaze detection techniques, the one used in the device used in the present study is Pupil Center Corneal Reflection (PCCR). The most modern eye-tracker devices are classified into the following types:

- Eye-tracker screen-base (or remote eye-tracker): they require a screen as a reference point for tracking the user's eyes;
- Eye-tracker wereable (eye-tracker glasses): these are wearable devices that are placed directly on the subject's body. They include devices such as goggles or other handheld devices that allow eye-tracking while the user interacts with his or her surroundings, without necessarily requiring a screen as a reference point.
- Integrated or embedded systems or augmented reality (AR) and extended reality (RX) devices.

A Tobii screen-based eye-tracker characterized by the use of infrared was used in this study, as the remote gaze detection system is based on the Pupil Center Corneal Reflection (PCCR) technique.

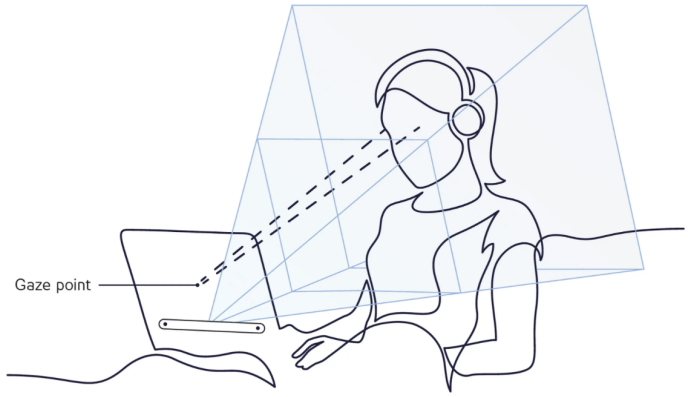


Figure 2.10: Example of eye-tracker screen-based

2.7.3.2 Pupillary and Corneal Reflexes Method (PCCR)

To estimate the fixation point, eye-trackers use a method called pupillary center corneal reflection (PCCR). This method is based on identifying and localizing the center of the pupil and the reflection of light on the cornea (known as glint). By means of a near-infrared light source, illumination is generated that creates a light reflection on the cornea and pupil Image processing algorithms identify the center of the pupil and the reflection of light on the cornea. When we move our eyes, the position of the iris and pupil changes with respect to the corneal reflection.

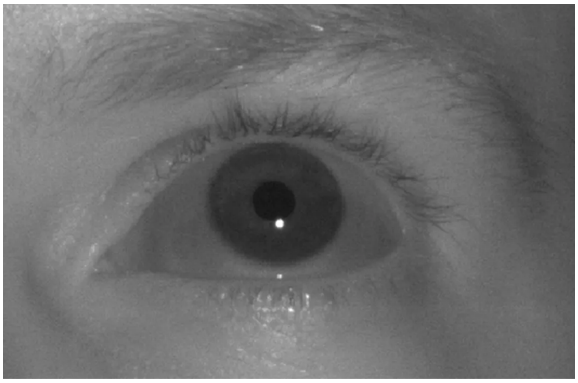


Figure 2.11: An image of an eye illuminated by the near-infrared light. The reflection of light on the cornea (glint) is tracked relative to the position of the pupil center. This allows the estimation of the direction of the point of gaze, or simply – what we look at.

Eye tracking algorithms use the vector between the center of the pupil and the corneal reflection and make further calculations to determine the position of the fixation point (i.e., what we are looking at) with great accuracy.

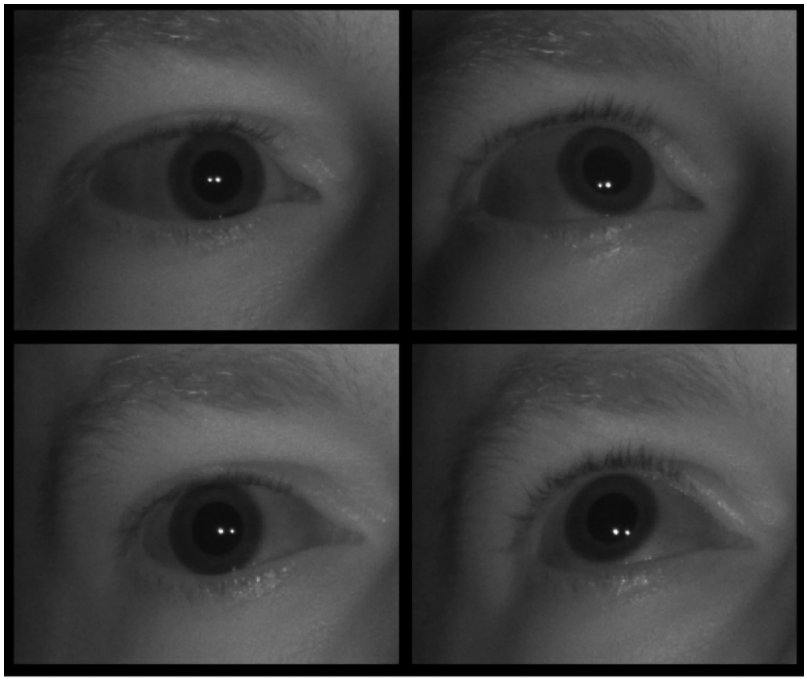


Figure 2.12: The position of the iris and the pupil changes with respect to the corneal reflection, which allows an accurate estimation of the point of gaze. Top images, from left to right: looking bottom left and top left. Bottom images, from left to right: looking bottom right and top right. Two glints are visible due to dual illumination

The accuracy of eye tracking depends on accurate annotation of the pupil center and corneal reflex. A stable illumination source that can withstand environmental variations and allow accurate annotation at any time is needed. Near-infrared light is an excellent light source that meets all the requirements for precise, high-accuracy eye tracking:

- Invisible to the human eye: near-infrared light is invisible to the human eye, therefore, it does not cause discomfort or distract the person in front of the eye tracker.
- Stable illumination under various conditions: compared with the visible spectrum, near-infrared light interferes less with ambient light. This ensures reliable and accurate illumination of the cornea, making eye tracking results stable in different environments [62].

Chapter 3

Method

3.1 Experimental Protocol

The research project has been submitted for evaluation by the Ethics Committee of the Neurology Unit of San Raffaele Hospital. A central element of the study is the NAP4A platform, an innovative tool used to assess cognitive and behavioral functions. This platform incorporates noninvasive, side-effect-free devices with the aim of early detection of signs of the preclinical stage of Alzheimer's disease. Subjects included in the study were selected following the inclusion criteria defined in the previous chapter. Before participating in the experimental phase, subjects undergo a neuropsychological evaluation with the aim of assessing cognitive domains. The scores obtained from the various tests are collected within a structured database, in which each column is associated with a defined parameter. In addition to the experimental data, the following are also indicated: subject's age, gender, level of schooling, and unique ID associated with the participant. Each participant's ID consists of the string "NAP4A" followed by a sequential, unique subject number and the letter "C" or "P" depending on whether the subject is a control or patient, respectively. The next phase constitutes the actual experimental phase and involves the acquisition of physiological signals using the sensors. The latter is divided into three distinct subphases:

- Basal phase (T0): represents the starting point of the study;
- First follow-up (T1): conducted after 6 months from the baseline phase;
- Second follow-up (T2): conducted after 12 from the start of the study.

Traditional neuropsychological assessment is administered only at baseline and at the second and final follow-up, so that any changes in cognitive function can be analyzed. Tests involving sensorimaging remain unchanged in all three phases, ensuring a consistent comparison over time. Two main activities are carried out during the experimental session:

- The first phase involves the administration, through the use of a tablet, of a battery of CANTAB digital cognitive tests. Through the performance of these tests, the assessment of daily activities, cognitive deficits, and behavioral characteristics is carried out, and physiological signals are acquired through wearable and non-wearable sensors. This phase involves the administration of the digital cognitive tests chosen from the CANTAB test battery; therefore, it is identified as the "Cantab phase."

- The second phase of the experimental session involves the administration of visual stimuli, which can be divided into two categories:
  - Personal images: photographs selected by the subject that evoke significant memories of his or her life;
  - Neffie images: standardized images used for experimental comparison.

  Given the presence of the Neuroaesthetic images, this phase is referred to as the "NEFFIE phase."

3.2 Cantab Phase Preparation

The first phase of the experiment involves the performance of a series of digital cognitive tests belonging to the CANTAB battery. During the performance, physiological signals are recorded through the use of biometric sensors to assess the subject's cognitive activity and emotional state. The sensors used in the "CANTAB phase" are: Shimmer GSR 3+, EEG helmet and Webcam. All these devices must be connected to the experimental platform before the start of the test by Bluetooth connection. Before each test, it is necessary to verify the connection of the sensors and the quality of contact with the subject so as to reduce artifacts during acquisition. Operationally, this session is performed on a tablet using CANTAB software. Before the start of the session, the neuropsychologist must log in and manually enter technical and subject data such as session date, participant's age, and ID code. In addition, preparation of the NAP4A platform prior to meeting the subject is expected by the back-end researcher. It is, in fact, necessary to set up six sessions for each participant, one for each selected test from the Cantab battery. The tests, in some cases, involve division into additional subtests shown in the table. The neuropsychologist during test execution and signal acquisition must press a random button, on a keyboard connected via USB connection, to record the time instants of the various tasks included in the same test. Battery execution follows a predetermined sequence of tests, each with a specific duration. The tests are administered in the following order:

- PRM1: duration of 2 minutes;
- MTT: duration of 4 minutes and 20 seconds;

- ERT: duration of 4 minutes;
- SWM: duration of 1 minute;
- PRM2: duration of 50 seconds;
- MTS: duration of 6 minutes;

Each battery test includes a preliminary instruction phase (during which instructions for performing the task are shown) and a training phase (a training session to allow the patient to become familiar with the task). Only after these two introductory phases does the test begin and data are saved for later analysis and evaluation of the cognitive domains. At the end of each test, there is a pause that allows the back-end researcher to save the acquisition just made and prepare the platform for the acquisition of the next test.

CANTAB TEST	**TASK**	**DURATION (sec)**
PRM1	EncodingI Immediate EncodingII	120
MTT	Direction Side Multitasking	260
ERT	All	240
SWM	four_cubes six_cubes eight_cubes	60
PRM2	All	50
MTS	All	360

Table 3.1: This table shows the duration (in seconds) of various tasks within the CANTAB test battery. Each row represents a specific task, with its corresponding subdivisions and execution times. Where "All" is indicated, only a single task is present.

Acquired signals are saved in zipped folders that include the presence of the excel file for

each signal, along with temporal and session information useful for later analysis. The folder name is always the same and includes the presence of the following information in the order given:

- Test phase (T0, T1 and T2)
- Cantab
- Test name
- Subject identification code followed by the letter "C" or "P" depending on whether the subject is a patient or control subject.

The structure of the zipped folder is as follows:

- Folder named "1 - new image 01", which contains as many .csv extension files as there are signals acquired:
  - EMOTIV -emotiv-, `EEG_PARAMETERS.csv`
  - EMOTIV -emotiv-,`EEG`
  - SHIMMER -gsr-, `GSR.csv`
  - SHIMMER -gsr-, `PPG.csv`
  - WEBCAM -webcam-, `FACE_EMOTION_5.csv`
- File `log.csv`, which contains information about the start and end timestamps of the test tasks;
- Files `playlist.csv`, `sensor-configuration.csv` and `session-info.csv`: files containing technical information about the recorded experimental phase.

3.3 Neffie Phase Preparation

The second phase involves a longer preparation and consists of the administration of visual stimuli. Each participant, at the first meeting, brings with him or her 16 personal images

imbued with emotional meaning and evoking particular moments in his or her life. These images are used during the test along with a set of standardized images called Neffie images. To ensure uniformity in the presentation of visual stimuli, all personal images are modified through the use of Adobe Photoshop by inserting a rectangular black border. This provides a consistent visual frame among all stimuli presented to subjects. All images administered during each subphase of the test, are identifiable by a unique code, and are catalogued and saved in an excel sheet specific to each participant. The file with the .xlsx extension is named with the subject's identification code and allows the images used in each session to be tracked. This phase of the experiment is divided into three main subphases:

- Training: this is useful to familiarize the subject with the way the test is conducted. During this phase, 4 images (2 personal images and 2 Neffie images) are shown to the participant.
- Encoding; the subject is subjected to more visual stimuli as 8 personal images and 8 Neffie images are used.

  In these first two steps, the image display has a duration of 11 seconds. Following the observation, there is the emotional rating step by means of a 3-second screen, during which the subject must verbally express the degree of liking of the displayed image. The level of liking is expressed on a scale from 1 (not very pleasant image) to 5 (very pleasant image). The neuropsychologist notes the scores on a table. Finally, there is the jitter phase during which the screen remains neutral for a duration of 6 seconds. This wait allows the participant to restore his or her attentional level. Following the encoding phase, to increase the time interval between encoding and recall, the neuropsychologist administers the ABAS II test to the participant, which lasts about 5 minutes. This is a pause to assess the subject's short-term memory during the next step.
- Recall: the last step of the experiment during which 28 images (14 personal images and 14 Neffie images) are administered. In this step, 8 images are the same images seen during the encoding phase and 6 are entirely new images. Then there will be an initial time interval, lasting 4 seconds, of viewing the image followed by a 3-second recognition phase during which the subject must express, in the form of a binary

response, whether he or she has already viewed the image in the previous step or not. The images are interspersed with a jitter period lasting 6 seconds.

The structure of the "Neffie phase" can be schematized by means of the following table:

TEST	TYPE OF STIMULUS	SUB-TEST	DURATION (sec)
Training	2 Neffie images + 2 personal images	Image Stimulus	11
		Raiting	3
		Jitter	6
Encoding	8 Neffie images + 8 personal images	Image Stimulus	11
		Raiting	3
		Jitter	6
Test ABAS			300
Recall	14 Neffie images and 14 personal images (6 from each group are the same as the encoding stimuli)	Image Stimulus	4
		Recognition	3
		Jitter	6

Table 3.2: Experimental Task Structure and Duration: this table outlines the different phases of the experiment, specifying the type of stimulus, sub-tests, and their respective durations in seconds.

Scores from the Encoding and Recall tests are structured in Excel sheets. During each subphase, there is the acquisition of signals by means of the same sensors as in the previous phase, with the addition of the eye-tracker that allows monitoring of eye movements and visual behavior of the subject during the administration of visual stimuli. The acquired signals are saved in zipped folders that include the presence of the excel file for each signal, along with temporal and session information useful for later analysis. The folder name is always the same and includes the presence of the following information in the order given:

- Test phase (T0, T1 and T2)

- Sub-test (Recall, Encoding and Training)
- Subject ID code followed by the letter “C” or “P” depending on whether the subject is a patient or control subject.

In the folder are files with the technical information of the experimental phase performed (“playlist.csv,” “sensor-configuration.csv,” and “session-info.csv”) and as many folders as there are images, raiting and jitter phases during the execution of the “Neffie” phase. Within each folder are the .csv files for each acquired signal:

- EMOTIV -emotiv-, `EEG_PARAMETERS.csv`
- EMOTIV -emotiv-, `EEG.csv`
- SHIMMER -gsr-, `GSR.csv`
- SHIMMER -gsr-, `PPG.csv`
- WEBCAM -webcam-, `FACE_EMOTION_5.csv`
- TOBII -tobii-, `GAZE`

3.4 Biomedical Signal Recording

Data acquisition is organized according to a precise division between back-end and front-end. From the back-end data, the protocol requires the setup and startup of the NAP4A technology platform on a PC, connected via USB cable to the webcam placed first on the tablet (during the “Cantab phase”) and later on the screen (during the “Neffie phase”). The PC is also connected to a keyboard, used by the neuropsychologist, who during the execution of the Cantab phase must press a causal key on the device to signal the beginning and ending time instants of each test task. The PRM1, MTT and SWM tests are divided into several sub-phases concerning different cognitive domains so each sub-phase must be marked by keyboard presses. These timestamps are saved in the.log file which are then fetched for signal pre-processing. The front-end configuration involves only the use of the tablet and webcam. When the participant arrives, the research team proceeds with

equipping the sensors: Shimmer GSR 3+ and EEG helmet. The GSR electrodes are placed on the index and middle fingers, and the optical pulse sensor is placed on the little finger. During acquisition, the participant is asked to reduce the movements of the hand on which the sensors are placed, which always corresponds to the non-dominant hand so as to reduce artifacts. The positioning of the EEG helmet requires special attention as it is highly influenced by the presence of glasses and hair, which reduce signal quality. It is always good practice to verify its correct positioning by feedback from the platform that provides the percentage of electrode contact with the subject's skin. After checking the correct connection of all devices, the neuropsychologist gives instructions and starts the Cantab test from the tablet. During this phase, the participant performs the required tasks, the back-end researcher observes the acquired signals in real time in order to intervene in case of any abnormalities. This experimental phase lasts longer than 20 minutes because there is a pause between tests to allow the researcher to properly start the next tests and reset the correct sensor placement. Next, the front-end is configured for the Neffie phase in which the use of the tablet is no longer necessary because the images are projected onto a screen connected via HDMI cable to the PC on which the NAP4A platform is installed. For this phase, the TOBII Eye-Tracker is installed on the bottom of the screen.

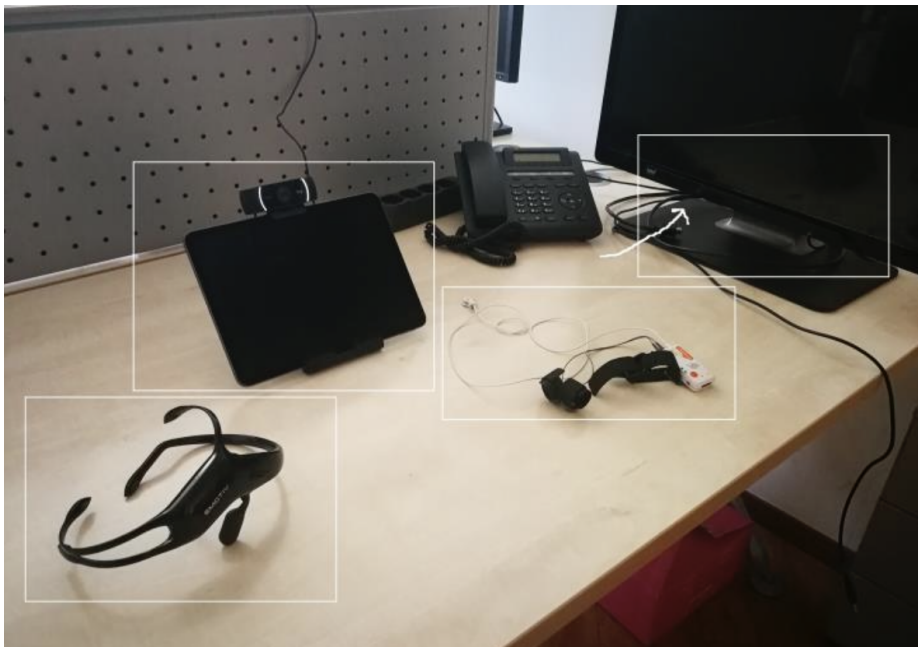


Figure 3.1: Assessment of sensor data acquisition during the experimental phase

At this point, the team proceeds with eye-tracker calibration and the monitor is duplicated to start the Neffie phase. The images are shown to the participant following a predetermined order articulated in the three sub-phases.

3.5 Database Building

The recorded data were saved in the MATLAB® environment through a multilevel structure. The saving of information was designed according to a hierarchical struct-based structure in order to organize the data systematically. The database has three macrostructures, each of which corresponds to a specific experimental phase, T0, T1 and T2. Within each phase, there are two additional subdivisions named C and P, to distinguish control subjects from patients, respectively. For each participant, a struct named *database_subjectcode* was created where the subject code is the unique code that identifies each individual participant in the experiment. This struct contains in turn, additional substructs related to the different sessions of the test and a binary variable that confirms the patient's actual participation in the experimental phase under consideration (T0, T1 and T2). Within the individual participant's struct, the data are divided into nine categories whose names correspond to the tasks performed during the "Cantab phase" (ERT, PRM1, PRM2, MTT, MTS and SWM) and the "Neffie phase" (Encoding, Recall, Training). If the task includes multiple subtests, each element has its own structure. Within the individual test structures there are two additional structs that provide binary information about the correct execution of the test and the presence of the file.log generated during the execution of the Cantab test in which the start and end time instants related to each task are indicated. In the case of Neffie tests, on the other hand, a struct is included indicating the presence of the playlist file, which is automatically generated during the experiment. During the performance of the various tests, biometric signals are acquired. The database includes these signals within the structs contained in each subtest structures named as the signal recorded (for example EEG, GSR, WEBCAM, PPG, TOBI). Here, signal are saved in a struct called "signal". In addition, the instants of time of signal acquisition saved in a struct named "time" are also recorded. Among the signals taken are: electroencephalography, photoplethysmography, skin conductance, and face registration for Cantab tests. In image-viewing tests, the eye-tracker signal is added to determine the subject's eye movements. Signal acquisition involves the generation of a file in .csv format in the patient's folder. Then, for immediate

verification of the presence of the data, a struct named "INFOSIG" is included in the database, which stores a binary information about the presence or absence of the file related to each signal. The structure related to signals includes a double-precision number matrix for the signal and a double-precision number matrix for the acquisition timestamp, except for TOBI data, as it is a coordinate pair. In addition to the recorded signal data, the database also includes details about the duration of each task and the timestamps for the start and end of each test. For the image display test, an additional struct stores in the form of a string, the code related to the image displayed by the subject.

3.6 Signal pre-processing

All signal analysis and pre-processing steps were performed in the MATLAB® environment. After organizing the data, pre-processing work was performed on the acquired signal. The pre-processing step of the "Cantab phase" involves an additional procedure compared to the Neffie phase because of the manual entry of timestamp flags during the Cantab tasks. Therefore, it was necessary to perform a check on the correct duration of the test to handle any delays or errors in the recording of times by the neuropsychologists. Since the expected length and the number of subphases present in each exercise are known, the timestamps present in the file.log are evaluated and, if necessary, the missing value is added by approximating the duration of the missed interval through appropriate function named *check_time2*. If the number of missing timestamps is greater than 2, dummy values are added to avoid errors. After checking, the total duration (expressed in milliseconds) of the task execution time is saved in the struct. Next, the sampling rate and its expected length were checked by implementing an algorithm. By evaluating the raw signals and the raw matrix of sampling timestamps, together with the start and end timestamps collected by the software in the excel file, both the expected and actual lengths are evaluated. By setting the sampling rate equal to 128 Hz, it is known that there should be a sample every 7.8125 ms (1/128). Knowing the first and last recorded timestamps, an ideal vector is created whose samples are spaced about 7.8125 ms apart. A cross-check is made by considering the value of the ideal vector just created and looking for a corresponding value in the acquired timestamps. If no value is found in the range of ± 5 ms from the ideal value, that sample is considered lost. After performing this check, the MATLAB® signal resampling function, using the cubic spline interpolation method, is applied to correctly resample the original

signal-time pair to the correct frequency as expected. The cubic spline interpolation method is a technique used to construct a curve through a series of data points, with the condition that the curve appears smooth. In cubic spline, each segment of the curve is described by a polynomial of degree three[64]. The output of this function consists of timestamp and signal vectors: if no missing values were found, these vectors are congruent with those originally sampled present and are saved in the "time" and "signal" structs. In the case where the algorithm has produced changes due to missing samples, the length of the original signal is compared with that of the expected signal. This procedure makes it possible to calculate the prediction rate of the signal as the difference between their lengths from the ideal length expressed by the following formulation:

$$\text{Percentage of prediction} = \frac{|\text{length ideal} - \text{length original}|}{\text{length ideal}} \times 100$$

At this point, each test or subtest provided by the protocol will contain a struct for each acquired signal, a struct containing information about the start and end timestamps, and a struct for the duration in milliseconds. Each signal, on the other hand, will have within it the original signal and timestamp vector, the predicted signal and its timestamp vector, and the information struct about the prediction percentage (value varying between 0 and 1). This data is critical to detect any sensor or platform malfunctions that affect subsequent feature extraction. In case of prediction percentage higher than the threshold value (set equal to 0.05%), the signals are not considered for further analysis. The next step is to assemble the entire signal for both experimental phases: the output of the "NEFFIE" phase is divided into a folder for each subphase for each administered image (image-rating-jitter). By concatenating all portions of the signal, the integer signal was reconstructed and saved in the struct "All." The same operation was performed for Cantab tests with multiple subtests. In this way each test includes the complete signal without differentiation between subtests. During the first acquisitions, an issue immediately emerged regarding the sampling rate of the electroencephalographic signal. The time-vector samples appeared to be repeated and were unevenly distributed, with differences between samples ranging from 1 to 15 milliseconds. For this reason, the signal was deemed unreliable, and the team began to solicit software developers to resolve the problem. Signal acquisition has been restored, but we do not currently have enough data to perform the analysis.

3.7 Eye-Tracker Signal

My thesis project mainly focused on the study of the eye-tracker signal with the aim of analyzing the distinguishing features between aMCI subject and control subjects. Despite the inconsistency of the database caused by the low number of patients recruited, extraction of specific features was carried out by me in order to identify significant patterns that can highlight the differences in eye behavior between the two groups. In the experimental protocol, eye-tracker signal acquisition is provided exclusively for image display tests (Training, Encoding and Recall).

3.7.1 Calibration of Eye-Tracker

The eye tracker must adapt algorithms to the front-end person in order to obtain gaze coordinate data with precision and accuracy. The adaptation process is called calibration during which the subject looks at points located at known coordinates. In the course of the research project, the calibration of the Tobii screen-based eye-tracker is initiated and controlled by the back-end researcher, who provides the subject with instructions for performing the process. Calibration of the eye-tracker is performed by showing the user a set of target points on the front screen, following the following steps:

- A small animated object is projected on the screen to attract the user's attention;
- All of a sudden, the object reaches the position of the calibration point and pauses for about 0.5 seconds so that the subject can focus on it. During the pause period, the size of the object decreases and the user's gaze is focused on the exact calibration point
  The animation in step 1 and the shrinking in step 2 should not be too fast because otherwise the result could be adversely affected, as you would not have enough time to focus on the target before the tracker starts collecting data.
- The tracker collects data for the specific point after the user has focused his or her gaze on the point, and a notification is sent to the application when the collection is complete;

- The next step is to move the object to the position of the next calibration point;
- All steps, so far listed, are repeated for all calibration points;
- The calibration result is calculated and displayed in a calibration graph;
- If the data for a given calibration point are missing or with low accuracy, the steps are repeated.

The normal number of calibration points is 2, 5 or 9. More can be used, but calibration quality does not improve significantly beyond 9 points. Usually, 5 points offer a very good result without being too invasive for the user. When choosing calibration points, it is important to consider the following aspects:

- Calibration points should cover an area equal to or greater than the area where the gaze-controlled application or stimuli will be shown, to ensure accurate data.
- Calibration points should be placed within the area tracked by the eye tracker and remain within the Active Viewing Area[65][66].

3.7.1.1 Eye-Gaze Estimation Algorithm

To identify the location of gaze on the screen to which the subject is focusing attention, mapping algorithms must be used. Gaze tracking algorithms include methods based on corneal reflection, which use infrared illumination (NIR) to estimate the direction of gaze or fixation point using polynomial functions or a geometric model of the human eye. Methods based on 2D regression and 3D model fall into this category. Before applying the gaze mapping algorithm, it is necessary to compensate for the effect of head movement[65]. The user's gaze position depends on the direction of gaze but also on the orientation of the head[67]. In PCCR techniques, if the user moves the head while always looking at the same point, the reflection vectors with respect to the pupil centers will be different from each other so the gaze positions will be inaccurate. The result of eye and head rotation is given by the following expression[65]:

$$dk - dk_{\text{ref}} = k(dk_{\text{gaze}} - dk_{\text{ref}})$$

Where dk_{ref} is the reference gaze direction of the user, dk is the head pose direction, and dk_{gaze} is the actual gaze direction (resulting from both eye and head rotation as shown in the 3.2). In the equation, k is a parameter related to head pan and tilt reported in [67] with values 0.5 and 0.4 respectively.

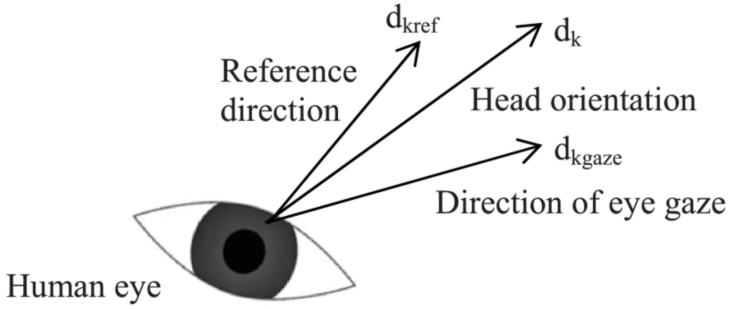


Figure 3.2: Relation between gaze direction and head pose[65].

3.7.1.1.1 2D Regression Based Methods

In regression-based methods, the vector between the center of the pupil and the corneal reflection is mapped to the corresponding gaze coordinates on the frontal screen using a polynomial transformation function. This mapping function is of the type f: (Xe, Ye) $\rightarrow$ (Xs, Ys) where Xe, Ye are the equipment coordinates and Xs , Ys are the screen coordinates. The relationship can be described by the following equations:

$$X_s = a_0 + \sum_{p=1}^{n} \sum_{i=0}^{p} a_{(i,p)} X_e^{(p-i)} Y_e^i$$

$$Y_s = b_0 + \sum_{p=1}^{n} \sum_{i=0}^{p} b_{(i,p)} X_e^{(p-i)} Y_e^i$$

Where n is the order of the polynomial and a_0 and b_0 are the coefficients. The polynomial is optimized through the calibration process during which the user is asked to fix certain

points on the front screen. The order and coefficients of the polynomial are selected so as to minimize the mean square difference (ε) between the estimated and actual screen coordinates expressed as:

$$\varepsilon = (X_S - M_a)^T (X_S - M_a) + (Y_S - M_b)^T (Y_S - M_b)$$

where a and b are the coefficient vectors and M is the transformation matrix[65].

3.7.1.1.2 3D Model Based Methods

With the three-dimensional method, the center of the cornea, optical and visual axes, and gaze coordinates are estimated using a geometric model of the human eye. 3D model-based techniques vary depending on the type of calibration performed and whether one or more cameras are used for acquisition. The 3D model-based method estimates line-of-sight (LOS) direction. The visual axis, which connects the fovea and the center of curvature of the cornea, is a straight line. The line connecting the center of curvature of the cornea and the center of the pupil is known as the optic axis. The visual axis, which differs from the optic axis, determines the direction of gaze. This deviation, which is about 5 degrees, is expressed in terms of the kappa angle. The 3.3 below shows optical and visual axes:

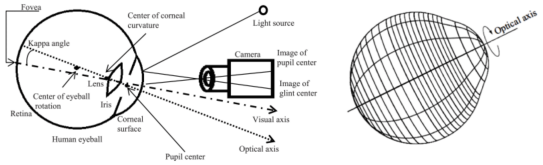


Figure 3.3: Model of a human eye ball, eye parameters and setup elements used in 3D eye gaze tracking. The optical axis is the line joining the center of curvature of the cornea with the pupil center. The visual axis passes through the fovea and the center of corneal curvature. Kappa angle is the angular deviation between the optical and visual axis.

When the three-dimensional positions of the cornea point and pupil center are known, it is possible to determine a subject's line of sight, which is represented by the direction of the visual axis. Therefore, the first step is to find out the coordinates of the center of the cornea.

To estimate this point, a stereoscopic system with two or more cameras and two infrared emitters can be used. The reflection laws of a convex mirror, simplified by the structure of the eyeball, state that an incident ray pointed in the direction of the center of curvature is reflected along the same path as the incident. The reflections of the incident rays are represented by the two reflected rays. Therefore, from these two lines can be drawn through the center of the cornea. The coordinates of the center of curvature of the cornea can then be estimated in three dimensions by intersecting the directions of the two incident infrared rays. Therefore, it is necessary to determine the three-dimensional position of the center of the pupil. Due to refractive phenomena at the interface between air humor and water, the image acquired by the camera represents the virtual projection of the pupil, which is located at a slightly extended position relative to the anatomical position of the pupil. Due to the symmetry of the pupil, this virtual projection is also located on the optical axis. Through the 3D coordinates of the pupil and those of the center of the cornea, an estimate of the optic axis can be obtained. Since the fovea is invisible from outside the eye, it is impossible to directly estimate the direction of the LOS. Through the use of the kappa angle, however, it is possible to estimate the rotation matrix that allows the direction of the visual axis to be obtained from the newly calculated optic axis management[65].

3.7.1.2 Eye- Tracker Data Quality

The quality of data acquired by eye-tracker is defined by three parameters: accuracy, precision and robustness. Regarding accuracy, a distinction is made: spatial accuracy and temporal accuracy. Spatial accuracy is defined as the difference between the coordinates of the actual position of the gazer and the position detected by the sensor. It is, usually, referred to as an error measured in degrees of field of view. Accuracy is considered a systematic error, and a higher value indicates lower accuracy. Temporal accuracy is defined by the time at which fixation events are reported. Poor temporal accuracy indicates the presence of delay between the onset of the eye fixation event and the recorded timpestamp. Accuracy, on the other hand, is a random error that indicates the repeatability of the gaze position detected by the eye-tracker when the actual position remains unchanged between trials. Accuracy also referred to as standard deviation. The difference between precision and accuracy is defined by the figure 3.4.

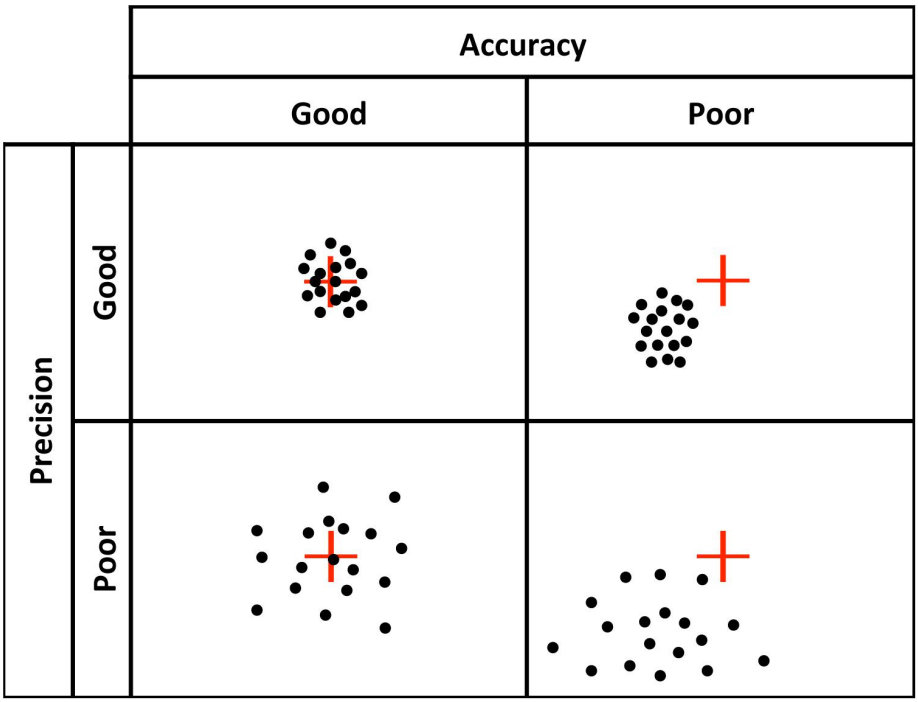


Figure 3.4: Visual representation of good vs. poor accuracy and precision[68].

Robustness refers to the amount of data lost as it is indicative of the amount of data recorded versus data lost during recording. Low spatial precision corresponds to noisier data (i.e., less signal) and can affect the output of fixation and saccade classification algorithms (Holmqvist et al., 2011; Blignaut and Wium, 2014), introducing artifacts that lead to shorter (Wass et al., 2013) or longer fixations (Holmqvist et al., 2012; Hessels et al., 2017)[68].

3.7.1.3 Eye-tracker Signal Filtering

The output of the eye-tracker consists of a two-column matrix containing the gaze coordinates (x,y) and a vector containing the time instants of the acquisition. These data provide information about the eye movement patterns. Below are plots of the eye-tracker signal for the control subject identified with the number 9:

- XY coordinates in the plan

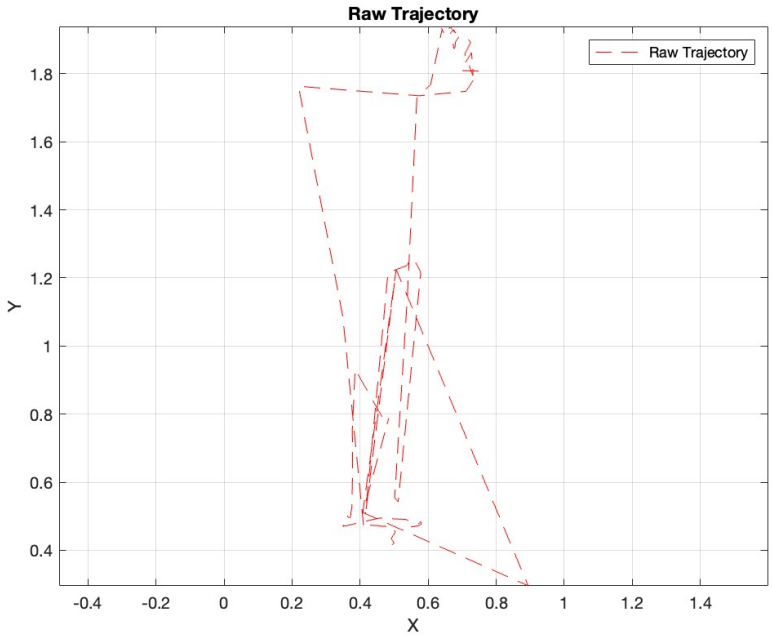


Figure 3.5: Trajectory of gaze movements in the XY plane.

- X and Y coordinates separately in time

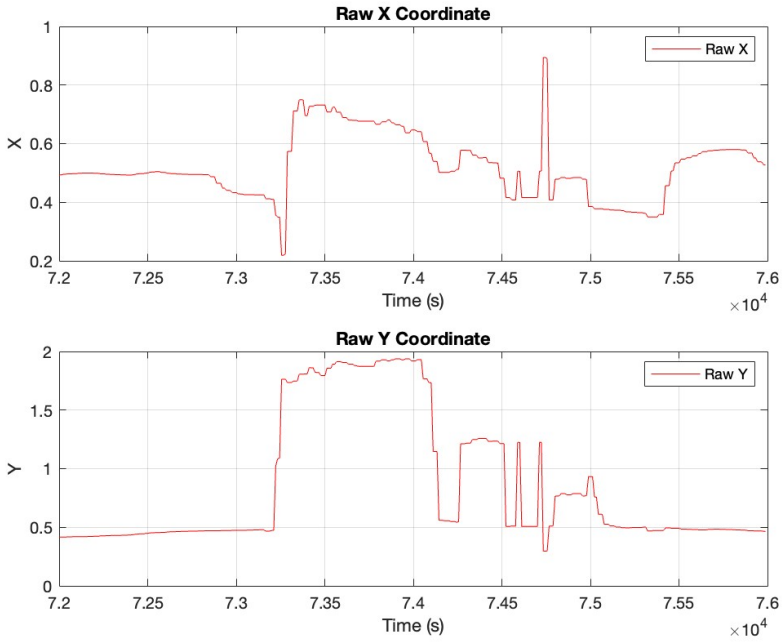


Figure 3.6: Evolution of gaze coordinates (X and Y) over time. Fixations appear as stable intervals, while saccades correspond to rapid changes.

Eye-tracker spatial and temporal resolution, participant, and artifacts affect signal quality[69]. Factors that can alter acquired signals include: blinks, microsaccades and rapid changes in fixation that increase the complexity of data analysis. Changes in ambient illumination can also affect signal quality. The first steps of signal processing consist of applying filtering and denoising techniques. The filtering step was performed by means of a special algorithm in the Matlab environment called *Tobi_filt*, which involves the application of a Savitzky-Golay filter (Savitzky & Golay, 1964). This filter makes it possible to preserve the characteristics of the high-frequency signal while maintaining its original shape and without distorting the velocity and acceleration information. It is a particularly suitable filtering for the eye-tracker signal as it performs local polynomial regression ensuring signal quality as demonstrated in the studies of Dai et al. (2017) and Nyström & Holmqvist (2010). In the work of Raju et al. (2023), several filters, including the Savitzky-Golay filter, were compared to evaluate their effectiveness in preserving signal and reducing noise in eye-tracking data. In this study, filters were evaluated on fixation segments of 256 samples, with a sampling rate of 90 Hz, such as that used in the TOBI_filt function. The results

showed that the Savitzky-Golay filter is effective in maintaining signal integrity during the filtering process[70]. Another study by Ko et al. (2016) used a Savitzky-Golay filter to smooth ocular motion signals, with the goal of preserving the original signal properties. This approach allowed accurate analysis of eye movements between saccades, highlighting the effectiveness of the filter in the context of eye-tracking. The function *Tobi_filt* applies a second-order Savitzky-Golay filter. The order of the filter was chosen referring to the study of Sarmad Al-gawwa[71] who recommends the order of the polynomial between 1 and 3. The smoothing time window was set equal to twice the minimum saccade duration, set at 0.03 seconds[69]. The smoothing process can be described by the following mathematical formulation:

$$S_j^* = \sum_{i=-m}^{m} C_i S_{j+i} \tag{3.1}$$

where:

- S is the original signal,
- S^* is the processed signal,
- C_i represents the smoothing coefficient for index i,
- N is the total number of points in the smoothing window, given by $N = 2m + 1$,
- m denotes the half-width of the smoothing window,
- j is the running index of the ordinate data in the original dataset[71].

By applying the Savitzky-Golay filter in the Matlab environment, the SG coefficients were obtained, which, given the raw coordinates (x,y), allowed the smoothed output of velocity and acceleration to be obtained separately by convolution. Next, the velocity and acceleration values were calculated by Euclidean distance multiplied by the sampling rate is divided by the angle expressed in pixels:

$$\text{vel} = \frac{\sqrt{\text{velX}^2 + \text{velY}^2}}{\text{angleInPixelsH}} \times \text{samplingFreq}$$
$$\text{acc} = \frac{\sqrt{\text{accX}^2 + \text{accY}^2}}{\text{angleInPixelsH}} \times \text{samplingFreq}^2$$

The following plots show the trend of velocity and acceleration before and after filtering:

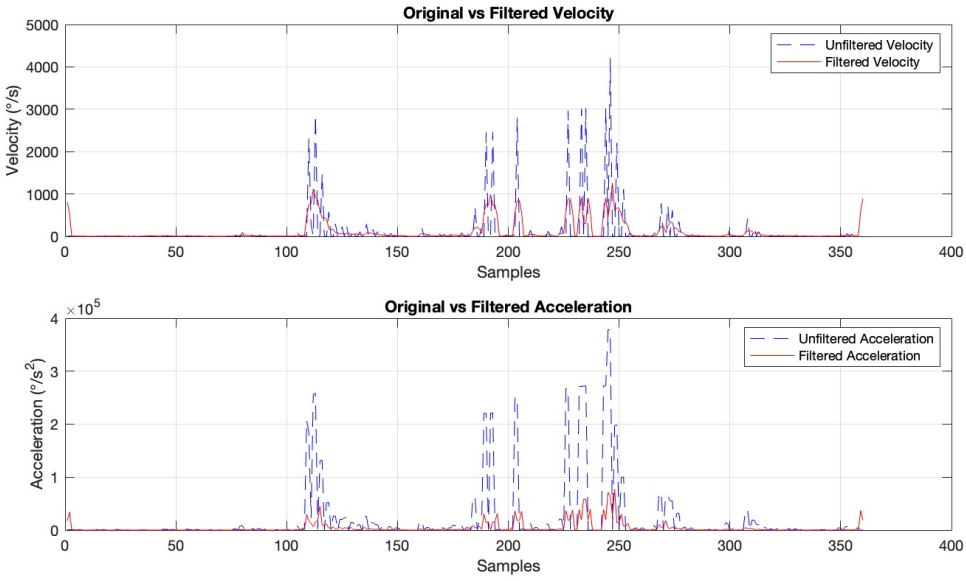


Figure 3.7: Unfiltered vs Filtered Velocity and Acceleration for Eye Tracking Data

To ensure reliability in the measurements, outliers exceeding the thresholds of 1000°/s and 100000°/s for velocity and acceleration, respectively, were excluded. These tresholds were identified based on the literature study by Duchowski, 2003 as higher values are associated with artifacts from blinking, instrumental noise, or acquisition errors. The signal trends before and after filtering are shown below:

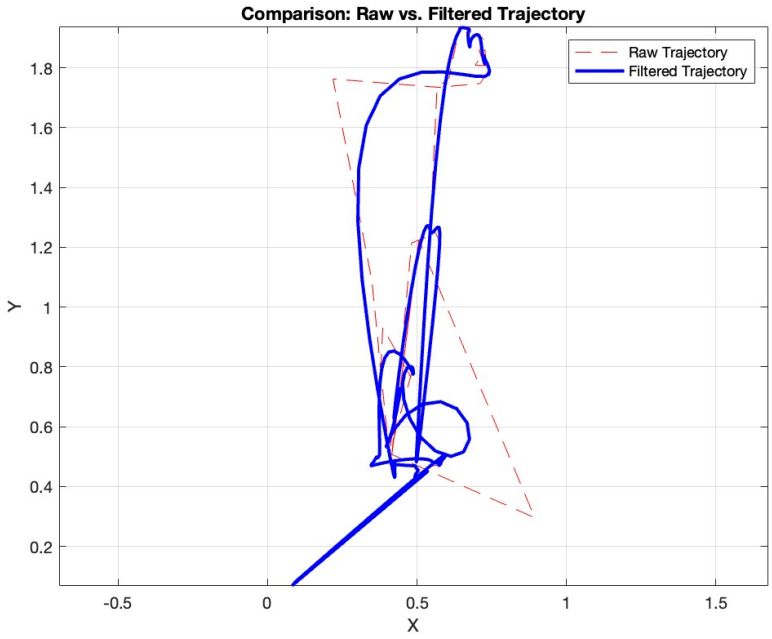


Figure 3.8: Comparison of the 2D trajectory in the X-Y plane: raw signal (red) vs. filtered signal (blue).

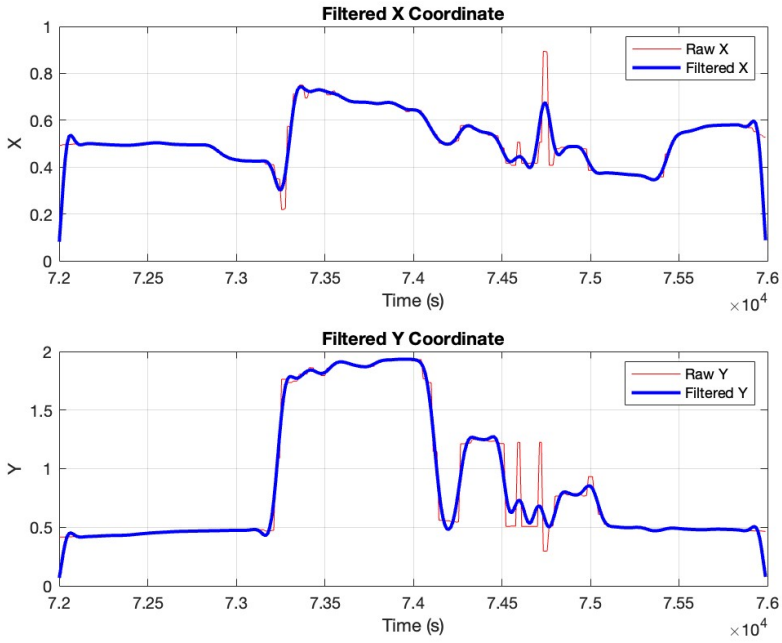


Figure 3.9: Comparison of the X and Y coordinates over time: the raw signal (red) and the filtered signal (blue).

The next function, implemented in the Matlab environment for noise removal is *Detect_NoiseTobi* which allow to handle blinks in position, velocity and acceleration data. The function takes as input the X and Y coordinates and the velocity threshold above which motion is considered a blink. Any abnormal data is detected and removed as it is considered as noise: if the percentage of sample affected by noise is greater than 20%, the function shows a *warning*. Otherwise, the function continues without interruption. After completing the filtering and denoising operations, the algorithm optimizes the threshold for peak velocity detection based on the noise during fixations. Peak velocity detection is performed by detecting peaks with velocity greater than the threshold. The choice of threshold value is critical because if it is too high, even the smallest saccades can be detected while if it is too small it could happen that samples belonging to a fixation are identified as belonging to a saccade. Initially the peak velocity detection threshold was set in the range of 100°/sec (according to the study by Marcus Nyström, 2010 [69] the threshold should be between 100°300°/sec). If the velocity is greater than the threshold, these are potenzial velocity

spikes or saccades. Samples that are not velocity spikes could be potential fixations, so they are labeled, using the function *bwlabel* as fixation groups in which the fixation duration is calculated. The minimum fixation duration was set to 100 ms, a value that is widely used to avoid micro-saccade inclusions. If the fixation duration is less than the minimum duration, the fixation is not considered. If, on the other hand, the fixation is long enough, the velocity data for the center of the fixation are extracted and the vector containing the velocity values of the fixations is updated. After processing all fixations, the mean and standard deviation of the vector containing the velocities of the fixations are calculated (denoted as avgNoise and stdNoise, respectively) For each iteration, the threshold for peak detection is updated following the trend: peakDetectionThreshold = avgNoise + 6 × stdNoise. The safety margin of 6 is considered a good robust level and is also used in microsaccade detection algo- rithms (Engbert & Kliegl, 2003). The threshold update is iterative, and the algorithm iterates until the difference between the current threshold and the previous threshold is less than a certain threshold (1°/sec). The spike detection threshold converges to a value such that saccades are detected. This algorithm updates according to the data and this ensures high robustness of the model. This method has two advantages because the choice of the threshold for pichi relection is data-dependent and not user-dependent, and, in addition, the threshold can be different for each user as it is adapted during the course of the algorithm. In addition, a threshold is defined for the saccade rate that is expressed according to the following expression[69]: saccadeVelocityThreshold = avgNoise + 3 × stdNoise Below is a table with a recap of the parameters set in the algorithm described so far:

Setting	**Value**
Filter	Savitzky–Golay (sgolay in MATLAB)
Filter order	2
Filter length	2xmin saccade duration
Peak velocity detection threshold	100°/sec
Max saccade velocity	1000°/sec
Max saccade acceleration	100,000°/sec^2
Min saccade duration	0.1 sec
Min fixation duration	0.03 sec

3.7.1.4 Detection of Eye Movement Events

From the literature analysis, variables were selected whose trends were visualized to enable validation of the eye-tracker as a diagnostic tool capable of identifying distinctions between the eye patterns of the control group and aMCI patients. The three starting parameters from which additional control variables were extracted are:

- fixations: fixations are identified as eye movements during which the retina is stabilized on an object for a certain period of time between 100 and 400 ms (as defined by the study [72]. The velocity during fixation is very low.
- saccades: saccades represent the period of time between two successive fixations. They are identified by rapid movements of both eyes used to reposition the central part of the eye to a new position in the visual environment. Saccadic movements, according to Duchowski, 2017, last between 10 and 100 ms.
- glissades: glissades indicate a disturbance in the final phase of the saccade during which the eye, instead of positioning itself at a fixed point, moves slowly past the point.

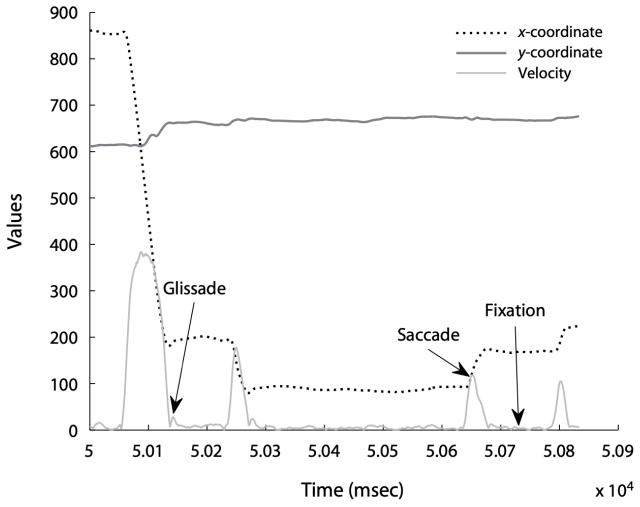


Figure 3.10: Example of an eye tracker signal with the main components identified: glissade (slow movements), fixation (visual pause), and saccade (rapid eye movements)

The following Matlab functions were implemented for parameter extraction:

- (*count_glissade*)
- (*count_fixation*)
- (*count_saccades*)
- (*detect_saccades_glissades*)

The function (*detect_saccades_glissades*) detects glissades and saccades. The inputs of the function are: velocity and acceleration vectors, Xand Y coordinates, and parameters that will determine thresholds for detecting saccades and glissades such as:

- minimum duration of fixation;
- minimum saccade duration;
- saccade velocity threshold;
- threshold for spike holding.

The function implements the following steps:

- Calculation of the mean and median of the eye movement velocity vector;
- Labeling of detected velocity spikes using the function (*bwlabel*);
- For each detected velocity spike, the function checks whether the duration of the probable detected saccade meets the minimum duration criteria to be considered a saccadic event, and if the velocity is sufficiently high then the saccade is considered as such;
- To detect the onset of the saccade, the first index is sought at which the velocity falls below the threshold and the acceleration is negative;
- Then a local threshold for velocity noise is defined by taking the mean and adding three standard deviations of the velocity before the onset of the saccade;

- The final instant of the saccade is sought;
- Finally, if the duration of the saccade is greater than the minimum threshold then all information about the saccade (amplitude, duration, beginning, end, and velocity) is saved. Otherwise it is excluded.

After identifying saccades, the algorithm looks for the presence of glissadic movements. Glissads are identified based on velocity thresholds that differ for weak glissads or strong glissads. All information regarding the glissade is saved in a vector. The algorithm used for the detection of glissades and saccades is the I-VT Algorithm (Identification by Velocity Threshold) based on the velocity threshold[4]. This method uses the velocities of eye movements to detect saccades. The velocity is calculated as the distance between the current point and the next point (it is measured in deg/sec). If the velocity exceeds the threshold of 300 deg/sec, the point is classified as saccade[73]. Next, by means of the function (*count_saccades*, saccade counting and analysis was performed. The presence of the number 1 in the binary vector, in output from the previous function, identifies the presence or absence of the saccade. This made it possible to count the number of saccades for each point in time (considering the sampling rate set at 90 Hz). The duration of the saccade was calculated by identifying its onset and offset, determined by the transition from 0 to 1 (or vice versa) in the vector, and then dividing by the sampling frequency. The distance traveled by the saccade is, in addition, calculated by the following Euclidean distance formula:

$$\sqrt{(\text{startPosX} - \text{endPosX})^2 + (\text{startPosY} - \text{endPosY})^2}.$$

Equally, the counting of glissades was carried out and their duration was derived by means of the function (*count_glissades*). Finally, the focus was on the study of fixations by means of the function (*count_fixation*). Fixations are temporal events during which the eye remains stable for a certain period of time defined by a threshold. An I-DT (Dispersion-Threshold Identification) algorithm was used throughout the study to identify and quantify fixations. This algorithm identifies fixations based on the dispersion between two consecutive points. The threshold for the minimum duration of fixations is set at 100 ms[72]. Two parameters must be set in this case: dispersion threshold and duration threshold. The duration threshold is generally between 100 and 200 ms, and in this case a value of 100 ms was selected. The dispersion threshold varies between 0.5-1° and in the course of the study

was set equal to 1°[73]. The dispersion algorithm uses a moving window, with an initial length equal to the duration threshold. In the selected window, dispersion is calculated by summing the difference between maxima and minima of the x and y coordinates according to the following formulation:

$$D = \frac{([\max(x) - \min(x)] + [\max(y) - \min(y)])}{2}$$

If the dispersion is greater than the dispersion threshold, the window does not represent a fixation and therefore the window moves one sample to the right. When the calculated dispersion is less than the dispersion threshold, then fixation is identified. To identify the fixation duration, the window is expanded until the dispersion exceeds the set threshold. The process continues by moving the window until the end of the length of the coordinate vectors.

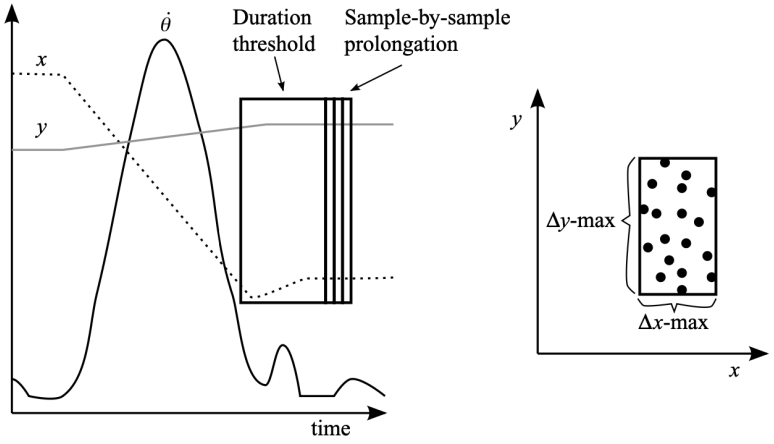


Figure 3.11: Diagram of the I-DT method for event detection, based on a duration threshold and sample-by-sample prolongation. On the right, representation of the maximum variation of parameters x and y for defining the segmentation criterion.[72]

The variables extracted from the three parameters described so far are listed in the following table:

Variable	Description
numFix	Number of fixations
meanFixDur	Mean fixation duration
medDurFix	Median fixation duration
maxFixDur	Maximum fixation duration
stdFixDur	Standard deviation of fixation duration
meanTimeNextFix	Mean time between successive fixations
medTimeNextFix	Median time between successive fixations
meanDistFix	Mean distance between fixations
stdDistFix	Standard deviation of the distance between fixations
latencyFirstFix	Latency of the first fixation
numSac	Number of saccades
meanSacDur	Mean saccade duration
meanSacDist	Mean saccade distance
maxDistSac	Maximum saccade distance
latencyFirstSac	Latency of the first saccade
numGlis	Number of glissades
meanGlisDur	Mean glissade duration
MeanAngVel	Mean angular velocity
MedianAngVel	Median angular velocity
visualEntropy	Visual entropy, which measures the variability or uncertainty of visual input during the observation period

Table 3.4: List of variables and their descriptions used in the analysis of eye-tracking data.

From these parameters, longitudinal analysis was performed to evaluate the evolution of eye patterns over time for the two groups (HC e aMCI) separately. In addition to the extracted variables with corresponding values, the output table contains the following information:

Field	Description
Subject ID	Unique identifier for the subject
Group	C (control) or P (experimental)
Experimental Phase	0, 1, 2
Cantab Column	Binary code (1 = cantab, 0 = not cantab)
Neffie Column	Binary code (1 = neffie, 0 = not neffie)
Test Type	Encoding, Training, Recall
Subtest Type	Jitter, Rating, Image Visualization
Image Code	Code of the image used for description

Table 3.5: Description of Experimental Variables

We only have the tests and subtests of the "Neffie phase" because the eye-tracker signal is only recorded in that phase.

3.8 Preliminary Statistic Analysis

3.8.1 Longitudinal Analysis of CANTAB Data

Preliminary statistical analysis allowed for a longitudinal analysis considering only those subjects who completed all experimental phases (T0, T1 and T2). Thus, the database subjected to statistical analysis includes 22 control subjects and 3 aMCI. The analysis was conducted on the results obtained from the Cantab test. The recruited subjects took the Cantab test to assess different cognitive functions. The parameters analyzed cover four main areas: Multitasking Test (MTT), Pattern Recognition Memory (PRM), Spatial Working Memory (SWM) and Emotion Recognition Task (ERT). For each area, the output variables were monitored at three different stages (T0, T1 and T2). By accessing the Cambridge University platform, it is possible to download the output in csv format containing all the results obtained from the digitized battery tests. Each row of the file represents a participating subject, while the columns contain the variables (about 150) along with additional information about age, gender, educational level, subject identification code, study phase (TO, T1, and T2), and experimental category (C or P). The team of

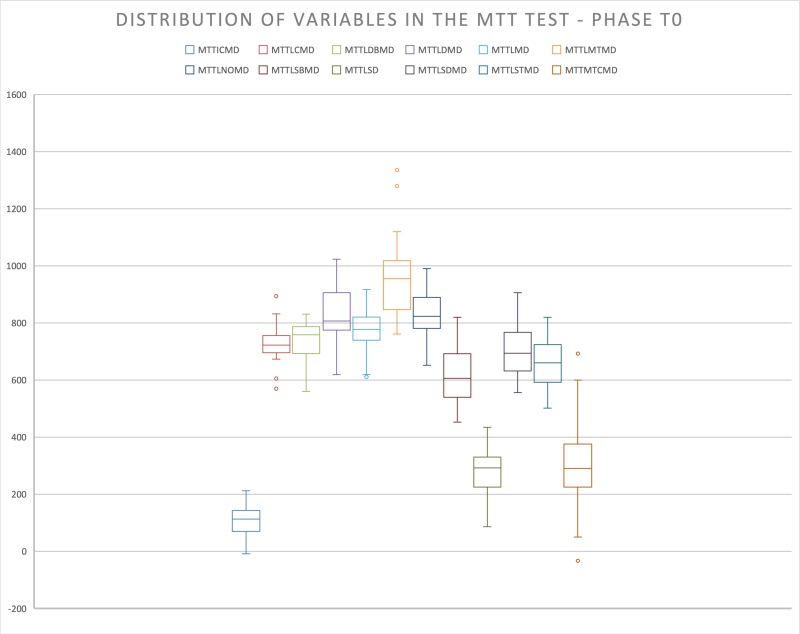
neuropsychologists contributed to the selections of the most significant variables for each Cantab test among all the variables in the output. The parameters analyzed are shown in the following table:

Cantab Variables	**Description**
Multitasking Test (MTT)	
MTTICMD - Incongruency Cost (Median)	The difference between the median latency of response (from stimulus appearance to button press) on congruent versus incongruent trials. A higher incongruency cost indicates that the subjects take longer to process conflicting information.
MTTLCMD - Median Congruent	The median latency of response on congruent trials in all assessed blocks.
MTTLDBMD - Median Direction Blocks	The mean latency of response in assessed block(s) in which the rule is to respond to the direction of the arrow.
MTTLDMD - Median Direction	The median latency of response in trials where the instruction was to respond to the direction of the stimulus.
MTTLMD - Reaction Latency (Median)	The median latency of response across all correct, assessed trials.
MTTLMTMD - Median-Multitasking Blocks	The median latency of response in assessed block(s) where both rules are used.
MTTLNOMD - Median Incongruent	The median latency of response on incongruent trials in all assessed blocks.
MTTLSBMD - Median Side Blocks	The median latency of response in assessed block(s) where the rule is to respond to the side of the screen.
MTTLSD - Dtandard deviation latency correct responses all trials	The standard deviation of the latency of response (from stimulus appearance to button press). Calculated across all correct, assessed trials.
MTTLSDMD - median latencyy side trials mixed	The median latency of response (from stimulus appearance to button press). Calculated for trials where the instruction was to respond to the side of the screen.

Cantab Variables	Description
MTTLSTMD - Median latency single task both direction and side	The median latency of response (from stimulus appearance to button press) in assessed block(s) in which only one rule is used.
MTTMTCMD - Median multitasking cost	The difference in median response latency between blocks using a single rule and blocks using both rules. Calculated by subtracting the median latency in single-task blocks from that in multitasking blocks. A positive score indicates slower responses during multitasking blocks.
Pattern Recognition Memory (PRM)	
PRMMDCLD - Median Correct Latency Delayed	The mean latency for a subject to correctly select the appropriate pattern during the delayed forced-choice condition, measured in milliseconds.
PRMMDCLI - Median Correct Latency Immediate	The median latency for a subject to correctly select the appropriate pattern during the immediate forced-choice condition, measured in milliseconds.
PRMPCD - Percent Correct Delayed	The number of correct patterns selected in the delayed forced-choice condition, expressed as a percentage.
Spatial Working Memory (SWM)	
SWMS - Strategy (6-8 Boxes)	The number of times a subject begins a new search pattern from the same box they started with previously. A low score indicates high strategy use.
Emotion Recognition Task (ERT)	
ERTCRT - Median Correct Reaction Time	The overall median latency for a subject to select the correct emotion based on the face stimulus.
ERTRT - Median Reaction Time	The median latency for a subject to select an emotion after being presented with a stimulus. Calculated across all assessed trials.
ERTTH - Total Hits	The total number of correct responses (emotion selection) across all assessed trials.

Table 3.6: List of Cantab variables and their descriptions in cognitive tests.

Given the small number of aMCI patients, a comparison analysis between the two groups was not performed but it was possible to observe the variation of variables separately for aMCIs and HCs. For control subjects (n=22), a longitudinal statistics method was applied, while for aMCIs, the trend of single-subject variables was displayed given the inconsistency of the sample. The first step for the healthy control subjects involved identifying outliers in the database using RStudio software, which allowed outliers to be detected within a range of $\pm$ 3 standard deviations. If the number of outliers was less than 20% of the data around the mean, the outliers were replaced with the median of the other values. The outliers were calculated separately for each experimental stage and for each variable considered, and their distribution is represented in the following boxplots.



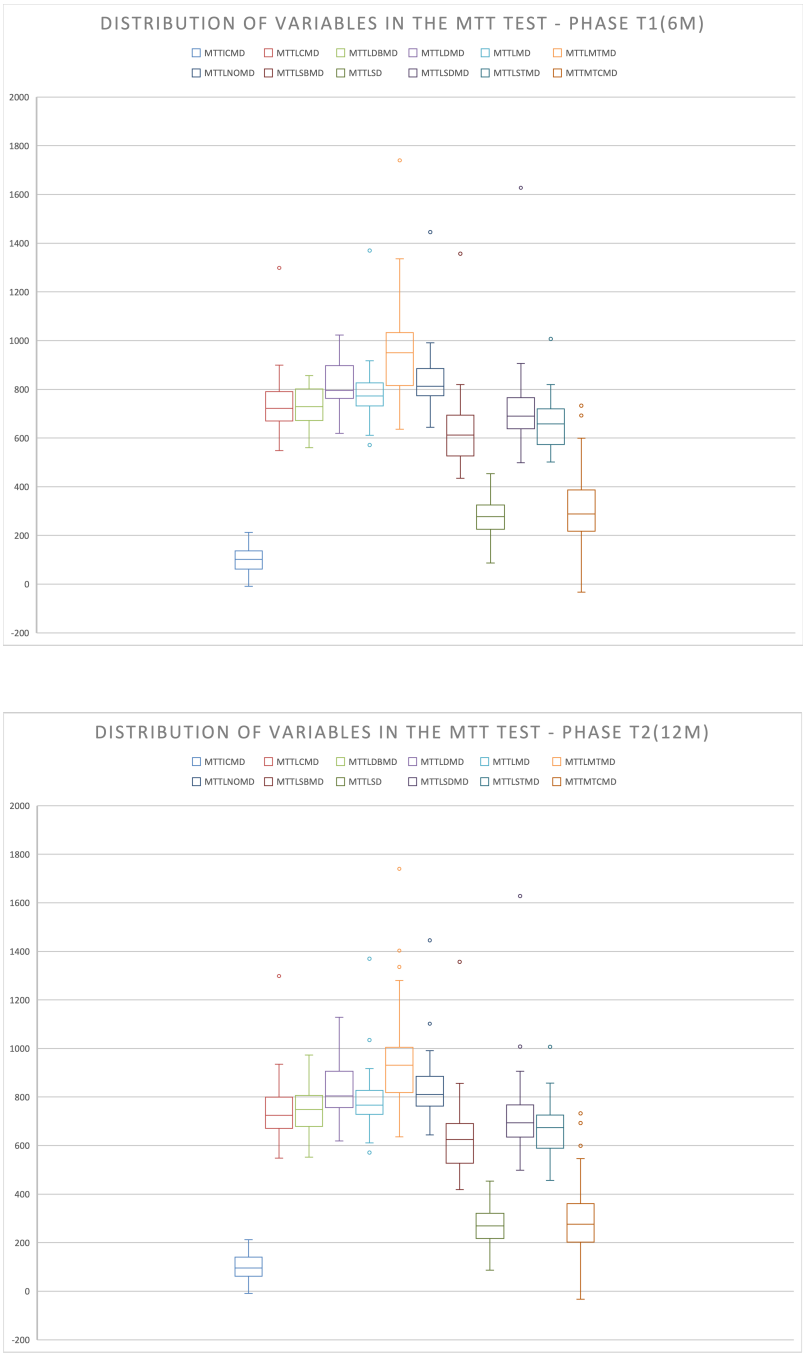


Figure 3.12: Distribution of variables in the MTT test across all phases (TO, T1, T2)

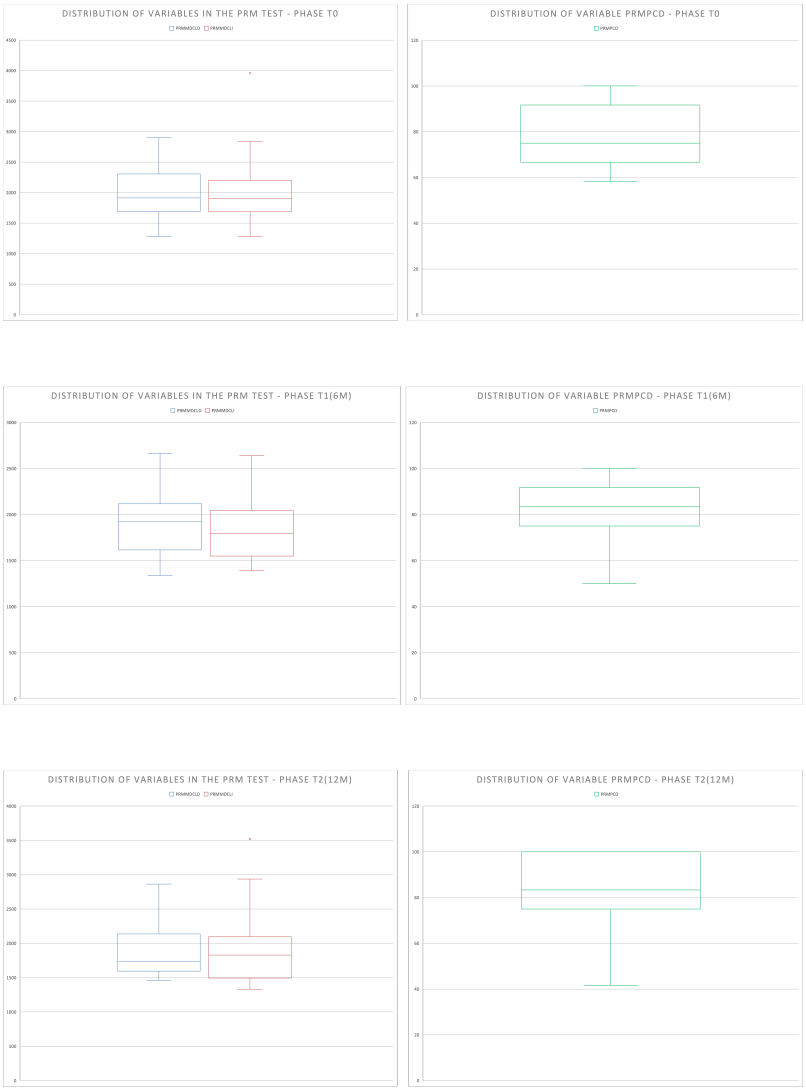
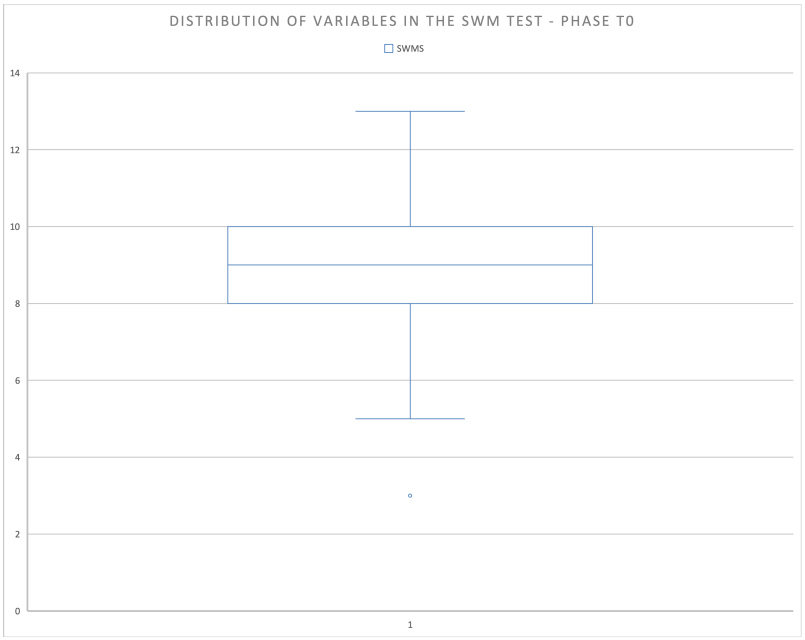
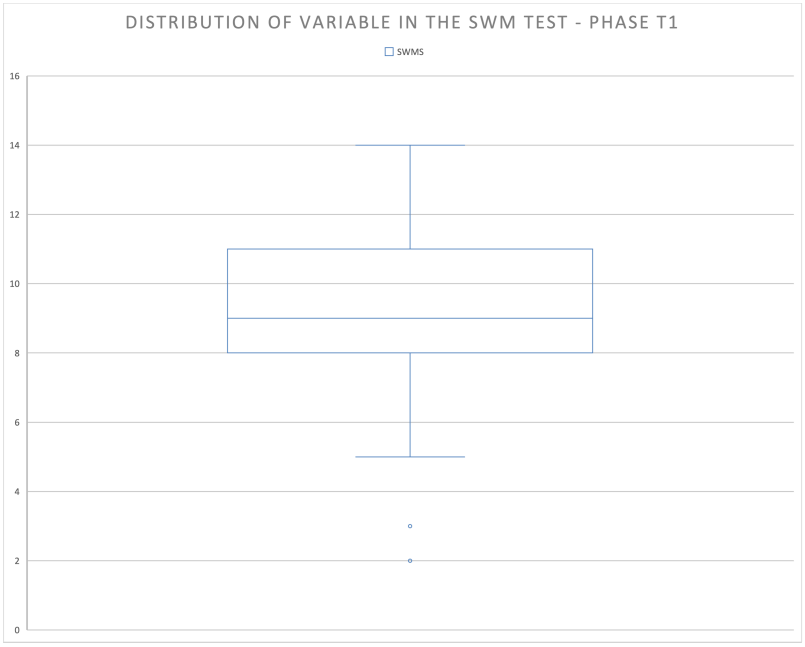


Figure 3.13: Distribution of variables in the PRM test across all phases (TO, T1, T2)


DISTRIBUTION OF VARIABLES IN THE SWM TEST - PHASE T0
SWMS
14
12
10
8
6
4
2
0
1


DISTRIBUTION OF VARIABLE IN THE SWM TEST - PHASE T1
SWMS
16
14
12
10
8
6
4
2
0

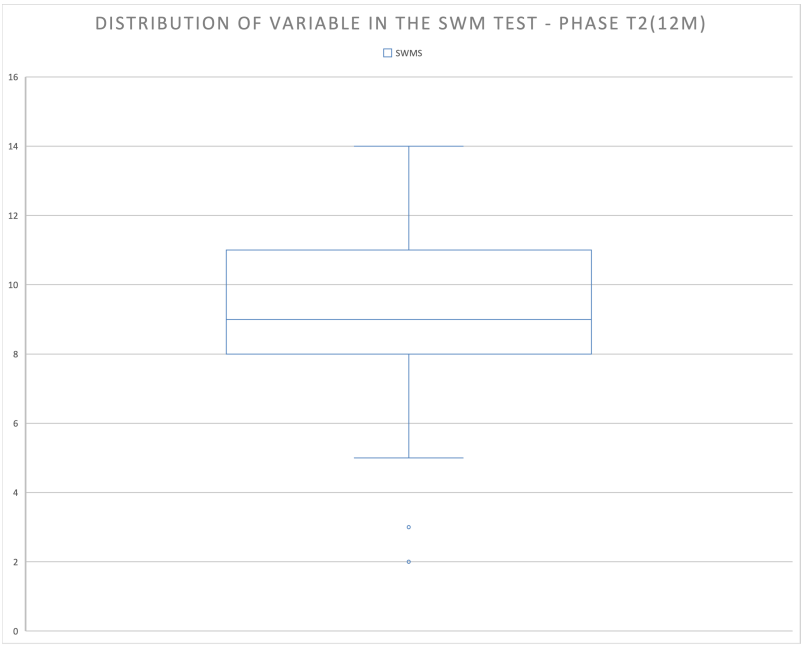


Figure 3.14: Distribution of variables in the SWM test across all phases (TO, T1, T2)

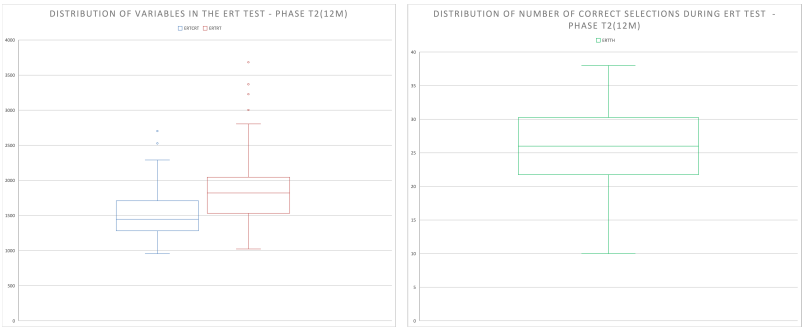


Figure 3.15: Distribution of variables in the ERT test across all phases (TO, T1, T2)

Next, the normality of the data was checked by means of the Shapiro-Wilk statistical test, selected because it is most suitable in the case of small sample sizes (< 50). The Shapiro-Wilk test for normality is given by the following formula:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

where:

- W is the value of the Shapiro-Wilk test,
- n is the number of observations in the sample,
- $x_{(i)}$ is the i-th ordered value of the data,
- a_i are the Shapiro-Wilk coefficients,
- $\bar{x}$ is the sample mean.

The parameter W takes values between 0 and 1. A value close to 1 identifies the normal distribution of the data[74]. After appropriate verification, longitudinal statistical analysis could be performed using the linear mixed-effects method. The evolution of the variables over time was performed by implementing the *lme4* library using RStudio software. A linear mixed-effects model was selected because it is suitable for longitudinal data because it allows conisdering:

- Random effects: take into account inter-subject variability (e.g., individual differences in response to the Cantab test);
- Fixed effects: independent variables of interest (e.g., sex, age, educational level)[75].

The model returns the *p-value* as output: if the value is below a conventional threshold (usually 0.05), it indicates a statistically significant difference between the values of the dependent variable at different stages of the study.

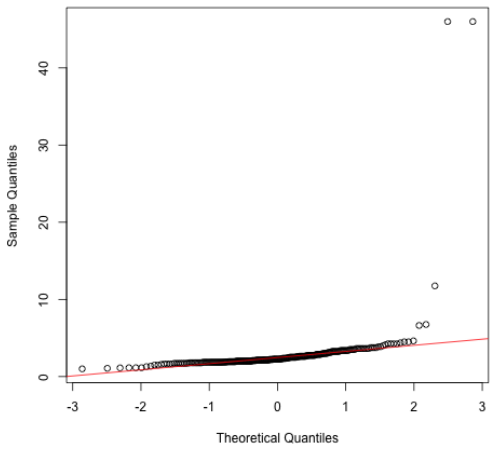
3.8.2 Analysis of eye-trakcer data

After implementing the algorithms and extracting the selected parameters described in Table 3.4, a statistical analysis was performed to identify eye patterns. Before proceeding with the analysis of the eye-tracker signal, a cleaning of the output table containing the variables extracted during the execution of the tests and subtests for each subject in the three experimental phases was performed. The cleaning made it possible to disregard the subphases of the Neffie phase for which the eye-tracker signal had a prediction percentage greater than 0.05% and the tests for which, a failure to synchronize the Tobi sensor, did not allow the signal to be recorded. In these cases, the number "-1" was present in the table, which was replaced with "NotANumber." In addition, by evaluating the expected eye position (provided by the calibration process) and the actual eye position, the "godness of fit (GoF)" indicator was also considered, which allowed for the elimination of items from unreliable measurements. The last step of output cleaning consists of removing outliers. From the structure of the "Neffie phase" it is known that three phases are planned to be performed: Training, Encoding and Recall but for statistical analysis only the Encoding and Recall phases were considered. In addition, for each phase there are the subphases of jitter, raiting and image display (personal or Neffie). For the analysis, features extracted only from the signals recorded during image display were evaluated. As with the Cantab tests, a statistical analysis comparing control subjects and aMCI patients was not outlined but studies were performed separately for the two groups. From Table 3.4, the most relevant parameters for the analysis of eye movement in aMCI established based on studies in the literature and support from the team of neuropsychologists at San Raffaele Hospital were selected. The analysis was conducted on the following parameters:

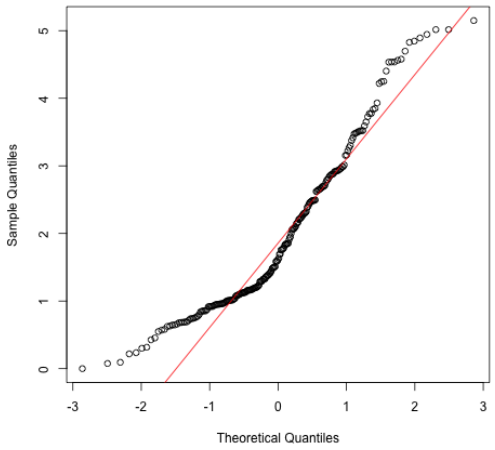
Parameter	Unit of Measurement
Number of fixations	Count (number)
Average fixation duration	Milliseconds (ms)
Visual entropy	Bits
Latency of first fixation	Milliseconds (ms)
Number of saccades	Count (number)
Duration of saccades	Milliseconds (ms)
Angular velocity of saccades	Degrees per second (°/s)

Table 3.7: Table of parameters with units of measurement

For the selected parameters, the mean, median and standard deviation values were calculated and organized in separate tables, where each row represents a subject and each column represents the mean, median and standard deviation of the analyzed feature. The information extracted from the analyzed signal parameters was divided into the case of Neffie visual stimuli and personal images, and in these two groups of stimuli, the Encoding and Recall stages were evaluated. To perform a preliminary analysis of the dataset, the first step was to assess normality in the distribution of the data considering all phases and all subjects in the control group. For this assessment, the Shapiro-Wilk test was performed, which demonstrated a nonnormal distribution in the data. The null hypothesis was rejected: this confirmed the need to separate the information for the different tests to better characterize the subject groups. The Q-Q plots of the examined features are shown below:

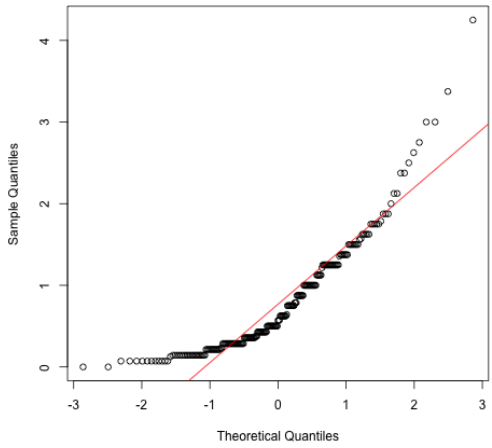


(a). Q-Q plot of the fixation count variable.

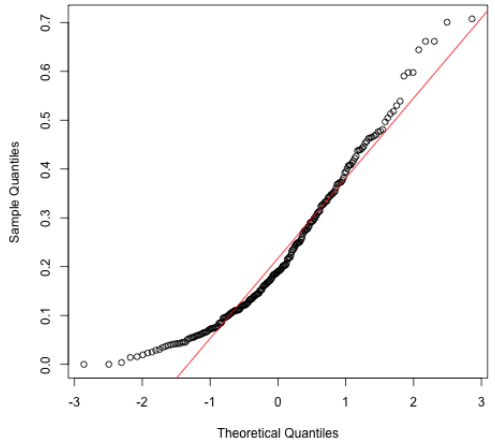


(b). Q-Q plot of the standard deviation of fixation duration variable.

Figure 3.16: Q-Q plots for fixation-related variables.

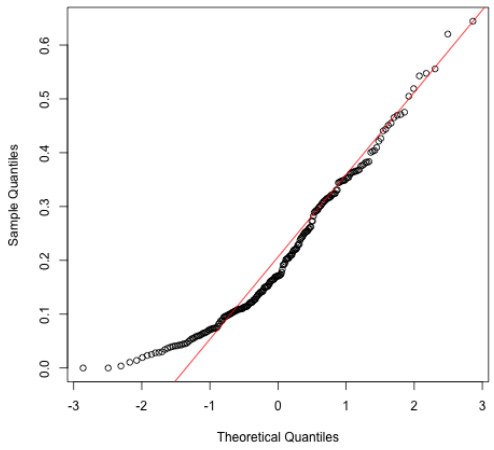


(a). Q-Q plot of the saccade count variable.

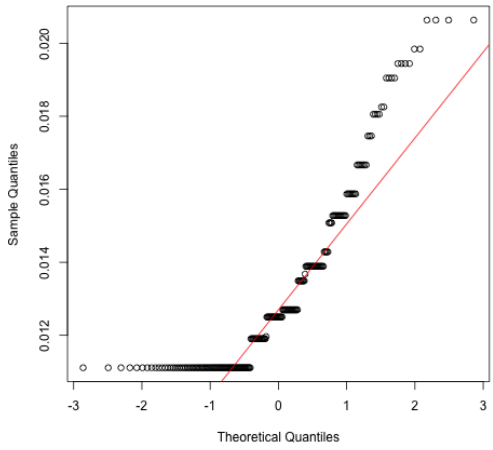


(b). Q-Q plot of the maximum saccade distance variable.

Figure 3.17: Q-Q plots for saccade-related variables.

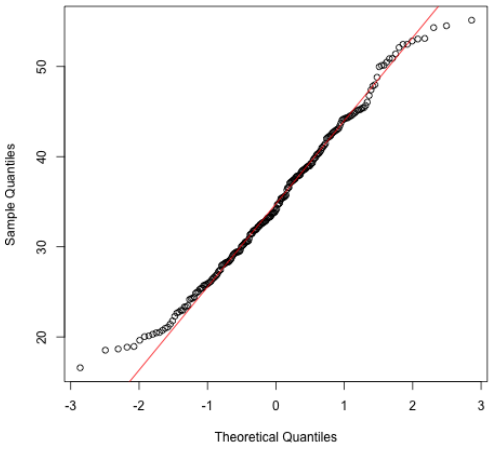


(a). Q-Q plot of the mean saccade distance variable.

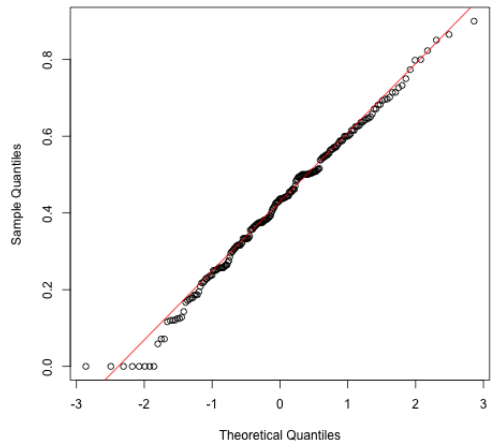


(b). Q-Q plot of the latency of the first fixation variable.

Figure 3.18: Q-Q plots for saccade distance and fixation latency variables.



(a). Q-Q plot of the mean angular velocity variable.



(b). Q-Q plot of the visual entropy variable.

Figure 3.19: Q-Q plots for angular velocity and visual entropy variables.

The non-normal distribution of the data required the use of a nonparametric statistical test to analyze the control group. The applied statistical test is the Wilcoxon-Mann-Whitney test, which is used to compare two independent data sets. In this case it is applied to compare

the same parameter, i.e., the extracted feature, between two different types of stimuli (Neffie and personal images) and between two different experimental phases (Encoding and Recall). This represents a first step for the goal of the NAP4A protocol since the experimental phases of the test and the types of stimuli go to investigate different domains. Encoding, for example, requires visualization of the image with subsequent expression of the level of pleasantness of the image: this requires a higher level of attention during the exploratory phase of the image while the Recall phase requires more mnemonic effort (as the subject is required to remember whether the image was viewed during the previous phase). The type of stimulus may also influence eye patterns in that personal imagery implies greater involvement resulting in greater visual attention and less eye movement. Although mean, median and standard deviation were calculated, the median was used for the statistical comparison because the Wilcoxon-Mann-Whitney test is more suitable for tests that do not follow a normal distribution and is less sensitive to outliers. This choice is supported by both the analysis of the distributions (Q-Q plot) and the existing literature in neurophysiology. All tests were performed using RStudio software, and the level of statistical significance was set at $p < 0.05$. In addition to p-value, rank statistic (to interpret the strength of difference between the two groups) and h-value (binary value indicating whether the hypothesis is null or not. If the p-value is less than 0.05 it returns a value of 1). This made it possible to identify how many and which variables show significant differences. In order to assess preliminary differences between control subjects and aMCI patients, a descriptive evaluation was performed. For the control group, the mean of participating subjects was evaluated for each phase and each stimulus evaluated. Then for a better visualization of the differences, graphs were used to display the scatter of the data where the straight line represents the mean value of the control group and the scatter points indicate the values of the three patients for the selected variable.

Chapter 4

Results

This chapter reports the results of the longitudinal analysis performed on the 22 control subjects and the single-subject descriptive analysis of the 3 aMCI patients who performed all experimental phases spaced 6 months apart. The socio-demographic characteristics of the recruited subjects are presented in Table 4.1: The gender ratio in each group was relatively

	aMCI (n = 3)	**HC (n = 22)**
Age	72.33 ± 11.55	65.91 ± 6.4
Education	16.67 ± 3.21	14.95 ± 4.81
Sex (F/M)	1 / 2	16 / 6

Table 4.1: Demographic and clinical characteristics of the sample. The table presents the distribution of participants by group (Healthy Controls - HC, and Amnestic Mild Cognitive Impairment - aMCI), sex (Male/Female), mean age ± standard deviation, and years of education ± standard deviation.

equal. In the HC group, 12 participants (54%) were male, and 10 participants (46%) were female. In the aMCI group, 2 patients (67%) were male, and 1 patient (33%) was female. The average age of the HC group was 68.5 ± 5.2 years, whereas the aMCI group had an average age of 70.1 ± 3.8 years. The minor age disparity between the two groups is deemed insignificant, as both groups were age-matched to mitigate possible age-related influences on cognitive performance. Regarding educational attainment, the HC group averaged 16.3 ± 2.5 years of education, while the aMCI group averaged 14.5 ± 2.8 years of education. Although this variation in years of education may affect cognitive performance, both groups possessed a relatively high educational level. These socio-demographic traits are crucial for understanding the context in which cognitive and clinical information is interpreted. Specifically, the age and educational background of participants must be considered when examining the cognitive performance of both the HC and aMCI groups, as these elements can influence cognitive skills and vulnerability to cognitive deterioration. However, to mitigate the effects of this independent variable, all results are corrected for education score.

4.1 Cantab Battery outcomes

The results obtained from the longitudinal analysis using a linear mixed-effects model on the control group are shown in Table 4.2. The variables are divided according to the

cognitive domain examined, and the mean, standard deviation, maximum, minimum and the significant level calculated in the three different experimental phases are given for each variable.

Cantab Variables	T0	T1	T2	*p-value*
MTTICMD - Incongruency Cost (Median)	107.07 ± 54.23	95.16 ± 43.61	91.55 ± 49.08	0.26
MTTLCMD - Median Congruent	725 ± 77.67	744.07 ± 151.37	731.18 ± 96.18	0.71
MTTLDBMD - Median Direction Blocks	732.43 ± 80.04	732.18 ± 82.96	747.2 ± 101.97	0.63
MTTLDMD - Median Direction	827.11 ± 95.15	817.66 ± 95.19	830.05 ± 114.39	0.76
MTTLMD - Reaction Latency (Median)	773 ± 79.24	790.55 ± 15.8	773.64 ± 101.12	0.75
MTTLMTMD - Median-Multitasking Blocks	962.36 ± 149.79	958.16 ± 223.93	926.91 ± 146.77	0.15
MTTLNOMD - Median Incongruent	832.07 ± 86.36	839.23 ± 158.3	822.73 ± 112.82	0.47
MTTLSBMD - Median Side Blocks	612.14 ± 95.51	634.91 ± 192.37	625.11 ± 111.36	0.32
MTTLSD - Reaction Latency (SD)	280.96 ± 77.03	272.19 ± 70.38	258.51 ± 59.31	0.24
MTTLSDMD - Median Side	703.66 ± 91.31	735.3 ± 220.14	706.73 ± 115.91	0.73
MTTLSTMD - Median Single Task Blocks	658.34 ± 81.94	664.02 ± 116.52	677.95 ± 101.91	0.49
MTTMTCMD - Multitasking Cost (Median)	304.02 ± 157.25	294.12 ± 153.16	248.95 ± 114.77	0.11
PRMMDCLD - Median Correct Latency Delayed	2011.8 ± 436.43	1917.32 ± 371.95	1892.75 ± 397.47	0.39
PRMMDCLI - Median Correct Latency Immediate	2015.86 ± 567.56	1877.39 ± 365.59	1902.86 ± 529.73	0.45
PRMPCD - Percent Correct Delayed	78.41 ± 13.53	81.82 ± 14.69	81.82 ± 16.19	0.48
SWMS - Strategy (6-8 Boxes)	8.73 ± 2.43	9.18 ± 3.25	8.27 ± 2.88	0.39
ERTCRT - Median Correct Reaction Time	1488.73 ± 419.89	1520.48 ± 343.2	1569.18 ± 378.62	0.54
ERTTH - Total Hits	25.91 ± 6.04	26.36 ± 5.52	25.59 ± 6.79	0.74
ERTRT - Median Reaction Time	1815.73 ± 599.16	1941.2 ± 506.78	1889.91 ± 458.22	0.05

Table 4.2: Values are means ± standard deviations. Statistical significance was accepted for values of $p < 0.05$

The application of the linear mixed-effects model, described in Chapter 3, made it possible to obtain the p-value for each variable evaluated in the three experimental phases. The p-value is a key parameter for interpreting the results. If the p-value is below a conventional threshold (usually 0.05), it means that there is a statistically significant difference between the values of the dependent variable in the different phases of the study. Not all Cantab variables have the same direction regarding value interpretation: some variables indicate a better performing test when they are higher while other variables are indicative of a less performing test when their value is higher. All selected variables, except ERTTH and PRMPCD, indicate a better performing test when they take lower values as defined by the Cantab test guide[57]. The following plots refer to the results obtained from the analysis

on RStudio for each Cantab test variable. The number "1" in the plots indicates that we have only one group under examination (healthy control group).

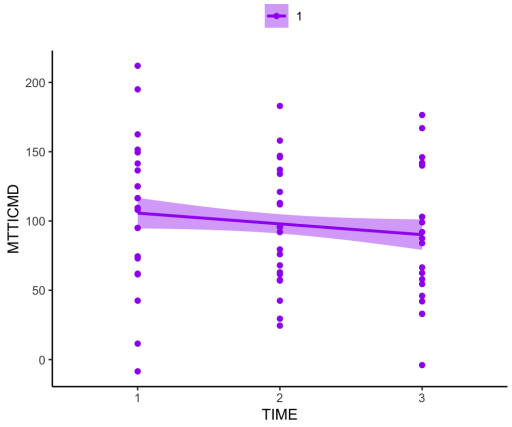


Figure 4.1: Trend of the variable MTTICMD (Incogruency Cost Median) during the three experimental phases.

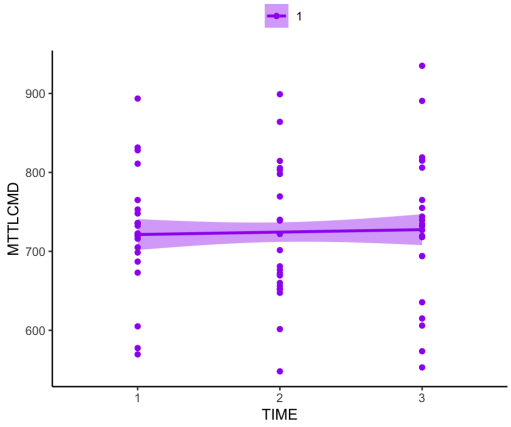


Figure 4.2: Trend of the variable MTTLCMD (Median latency - Congruent) during the three experimental phases.

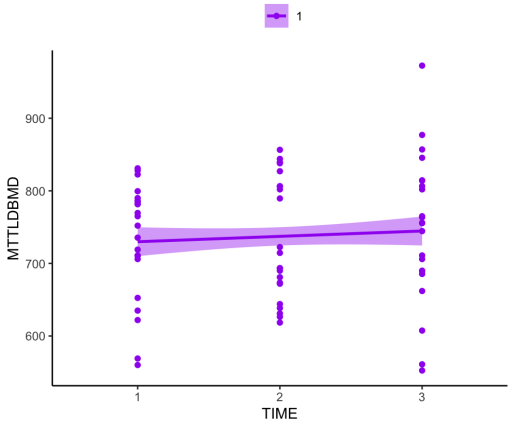


Figure 4.3: Trend of the variable MTTLDBMD (Mean latency for Direction Blocks) during the three experimental phases.

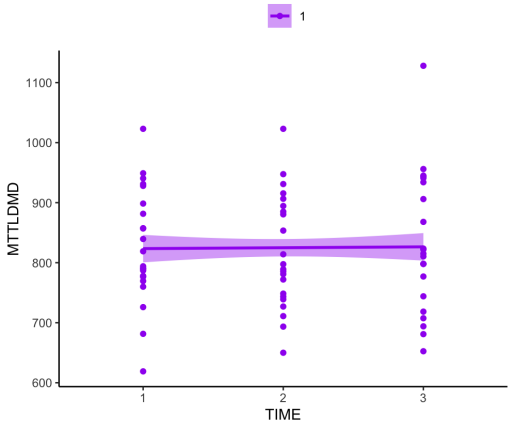


Figure 4.4: Trend of the variable MTTLDMD (Median latency Direction) during the three experimental phases.

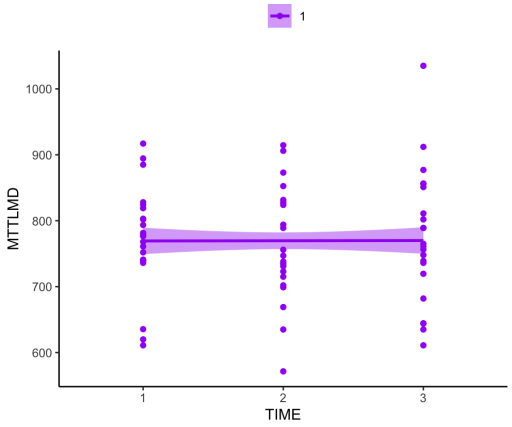


Figure 4.5: Trend of the variable MTTLMD (Reaction Latency) during the three experimental phases.

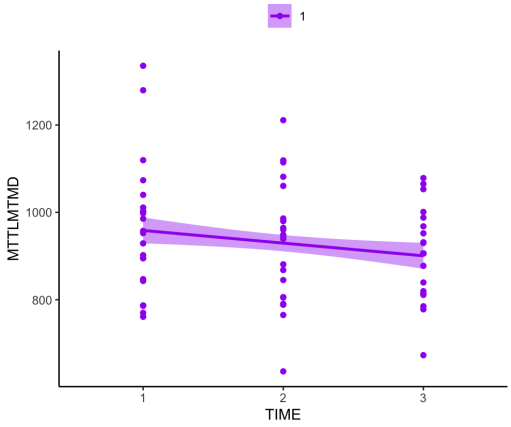


Figure 4.6: Trend of the variable MTTLMTMD (Median Multitasking Blocks) during the three experimental phases.

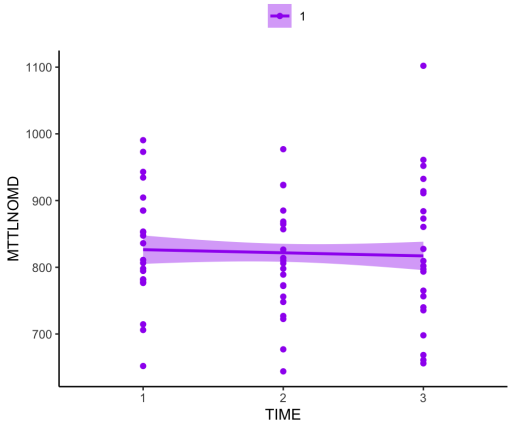


Figure 4.7: Trend of the variable MTTLNOMD (Median Incongruent) during the three experimental phases.

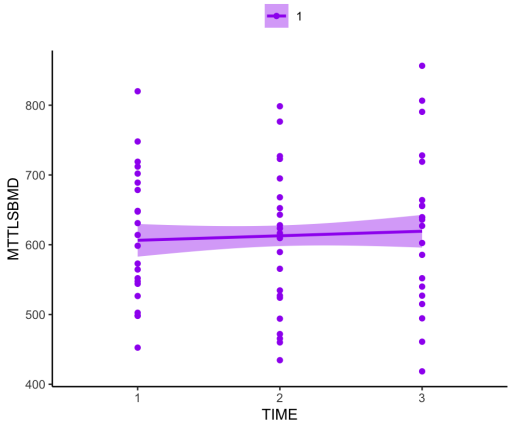


Figure 4.8: Trend of the variable MTTLSBMD (Median Side Blocks) during the three experimental phases.

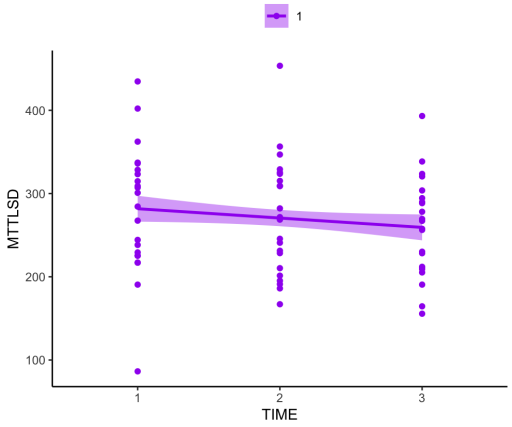


Figure 4.9: Trend of the variable MTTLSD (Standard Deviation latency correct responses all trials) during the three experimental phases.

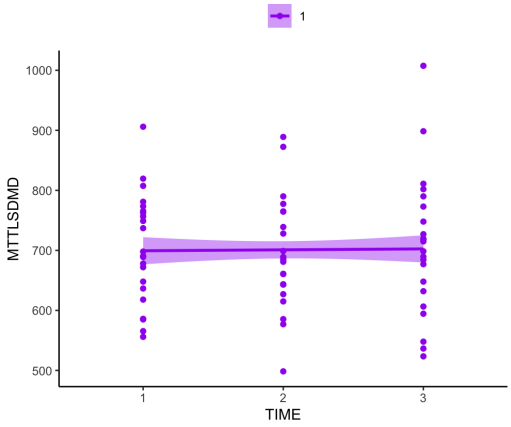


Figure 4.10: Trend of the variable MTTLSDMD (Median latency side trials mixed) during the three experimental phases.

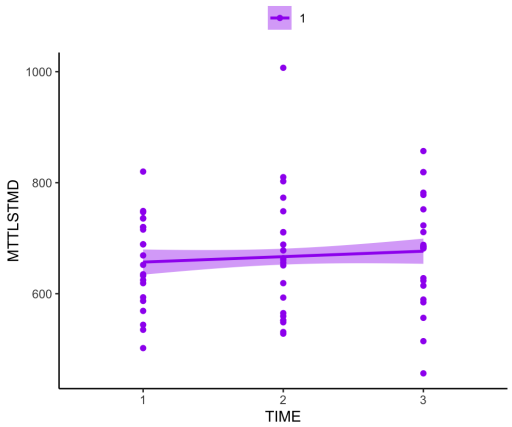


Figure 4.11: Trend of the variable MTTLSTMD (Median latency single task both direction and side) during the three experimental phases.

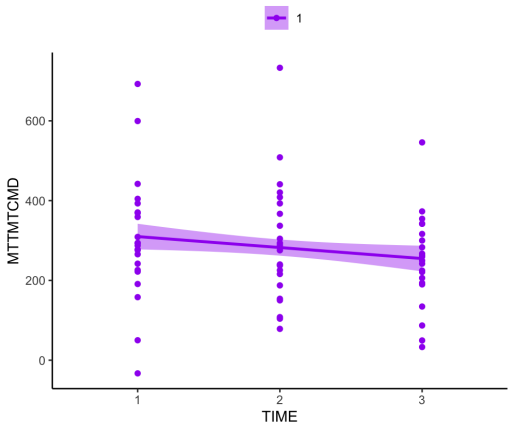


Figure 4.12: Trend of the variable MTTMTCMD (Median Multitasking cost) during the three experimental phases.

The Multitasking Test (MTT) was used to investigate the dimensions of multitasking by different parameters. The MTTICMD parameter did not show significant changes over time (p-*value*=0.26) indicating stability during the observation period. Similarly, other variables

such as MTTLDBMD, MTTLDMD and MTTLDMD showed *p-value* greater than 0.05 emphasizing stability in terms of latency, and direction. The MTTMTCMD, MTTLMTMD, MTTLSD tests showed a slight decrease, indicating better performance, but not significant.

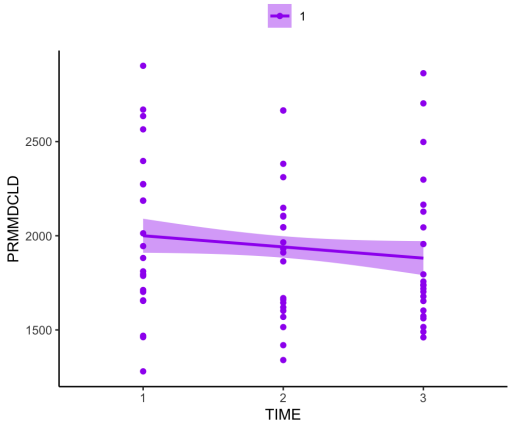


Figure 4.13: Trend of the variable PRMMDCLD (Median Correct latency delayed) during the three experimental phases.

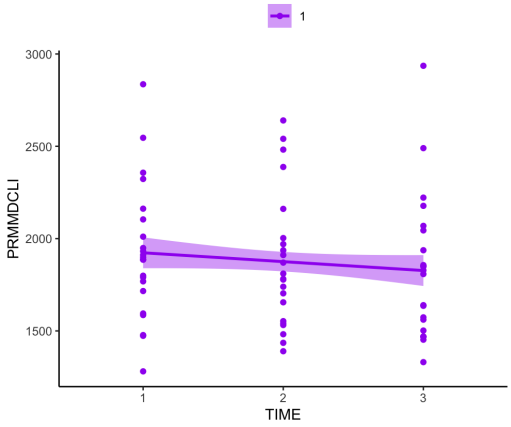


Figure 4.14: Trend of the variable PRMMDCLI (Median Correct Latency Immediate) during the three experimental phases.

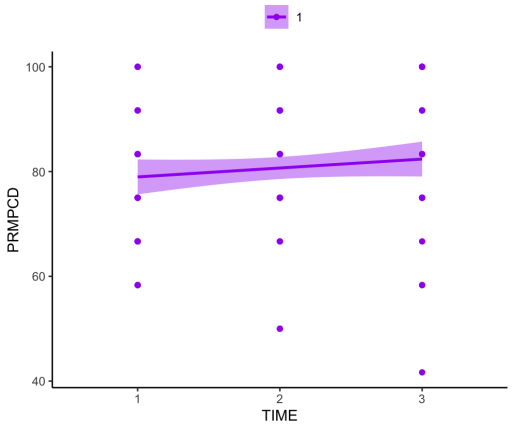


Figure 4.15: Trend of the variable PRMPCD (Percent Correct Delayed) during the three experimental phases.

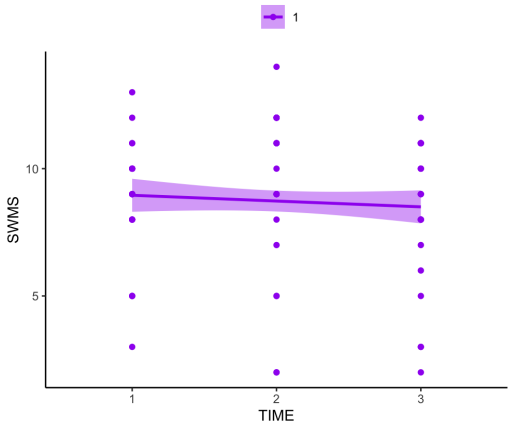


Figure 4.16: Trend of the variable SWMS (Strategy) during the three experimental phases.

Regarding the Pattern Recognition Memory (PRM) domain, the variables measuring the reaction time trend (PRMMDCLD and PRMMDCLI) showed no significant change (*p-value*>0.05) despite a slight decrease over the three experimental phases. The percentage of correct responses (PRMPCD) showed an increase that was not statistically significant. In the Spatial Working Memory tasks, the SWMS variable showed no significant change.

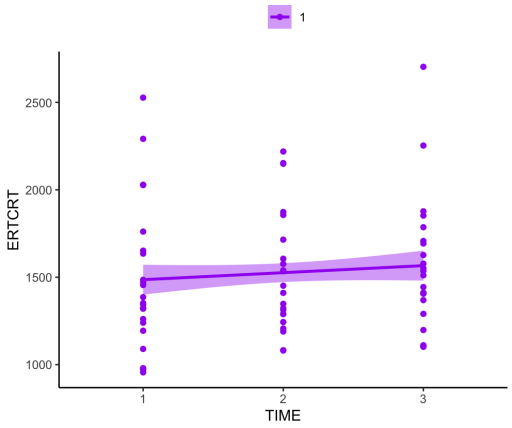


Figure 4.17: Trend of the variable ERTCRT (Median Correct Reaction Time) during the three experimental phases.

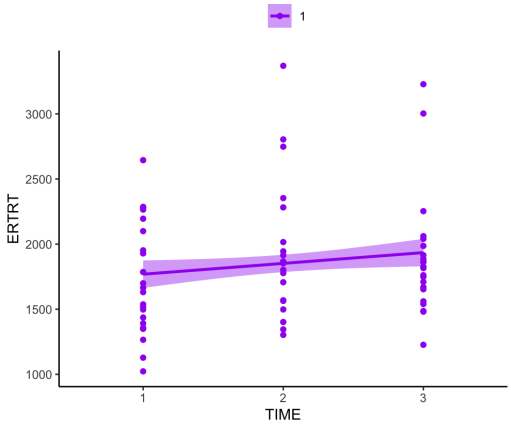


Figure 4.18: Trend of the variable ERTRT (Median Reaction Time) during the three experimental phases.

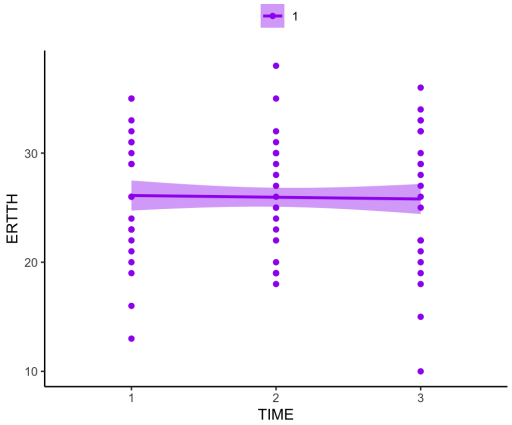


Figure 4.19: Trend of the variable ERTTH (Total Hits) during the three experimental phases.

Finally for the Emotion Recognition Task (ERT) test, the rather stable trend of reaction time variables (ERTRT, ERTCRT) was displayed. The variable indicating the percentage of correct responses (ERTTH) shows positive changes, not statistically significant. No variables, as expected from a control group, showed significant changes.

For aMCI patients, the trends of the variables for each test listed in Table 3.6 in the three experimental phases (T0, T1 and T2) were evaluated for individual subjects who are identified with following code: 8P, 10P and 18P. The letter "P" identify the patients group. In the individual plots, each curve represents data for one patient, with the values of the variable plotted as a function of the experimental phases.

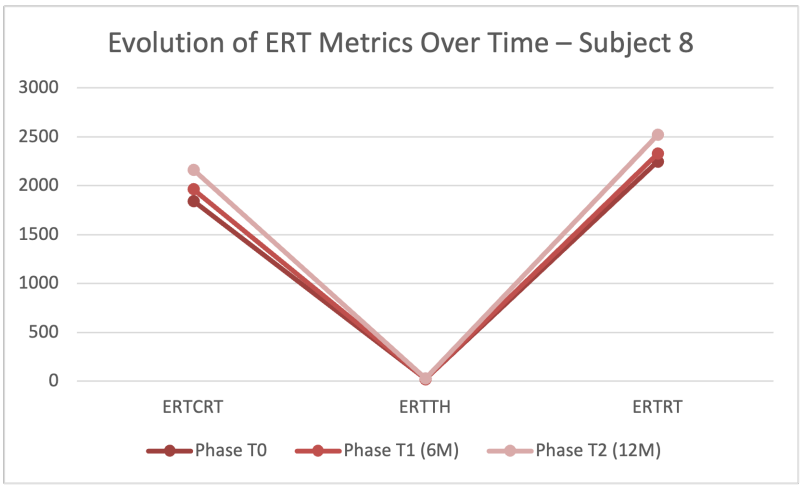


Figure 4.20: Evaluation of the variables ERTCRT, ERTTH and ERTRT during ERT test for the patient number 8P

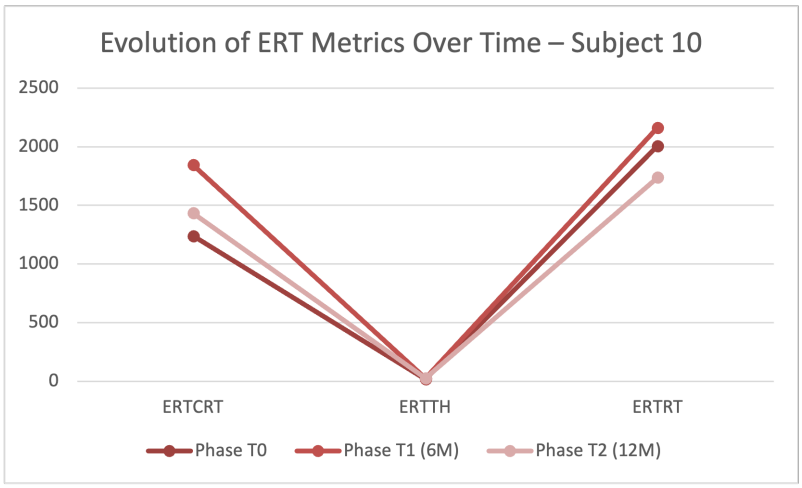


Figure 4.21: Evaluation of the variables ERTCRT, ERTTH and ERTRT during ERT test for the patient number 10P

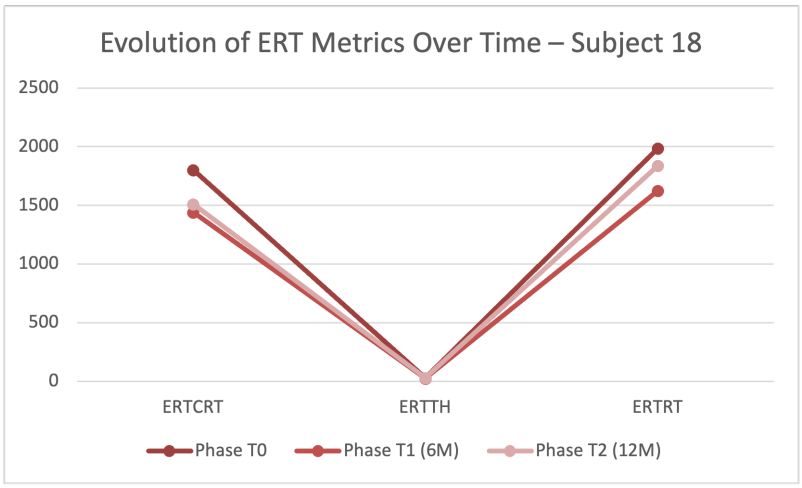


Figure 4.22: Evaluation of the variables ERTCRT, ERTTH and ERTRT during ERT test for the patient number 18P

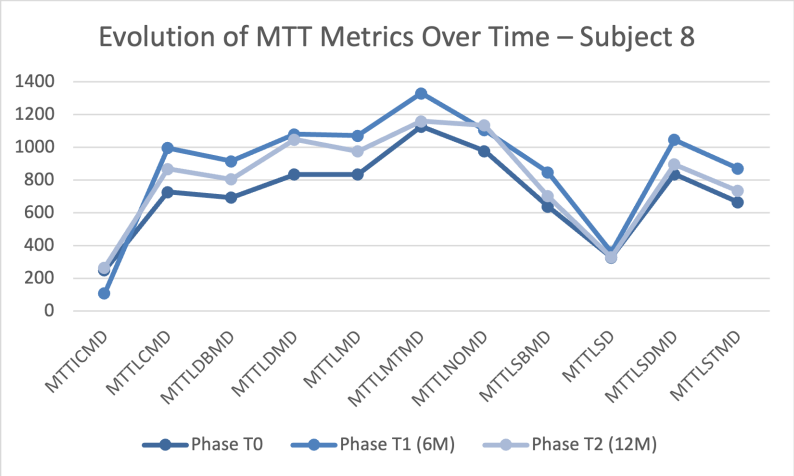


Figure 4.23: Evaluation of the variables MTTICMD, MTTLCMD, MTTLDBMD, MTTLDMD, MTTLMTMD, MTTLNOMD, MTTLSBMD, MTTLSD, MTTLSDMD, MTTLSTMD and MTTMTCMD during MTT test for the patient number 8P

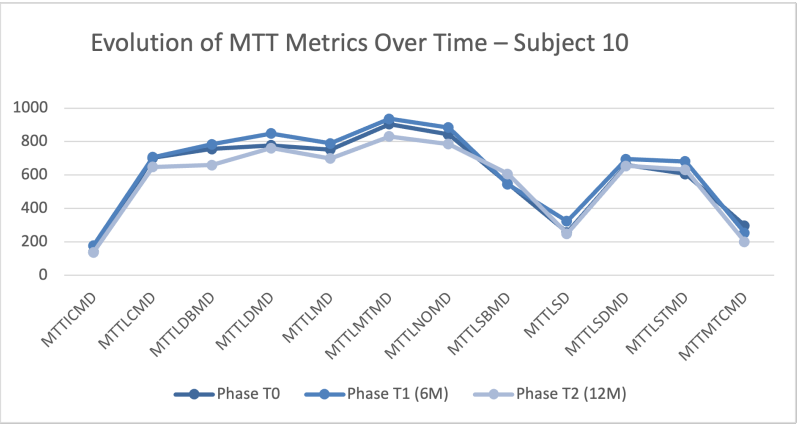


Figure 4.24: Evaluation of the variables MTTICMD, MTTLCMD, MTTLDBMD, MTTLDMD, MTTLMTMD, MTTLNOMD, MTTLSBMD, MTTLSD, MTTLSDMD, MTTLSTMD and MTTMTCMD during MTT test for the patient number 10P

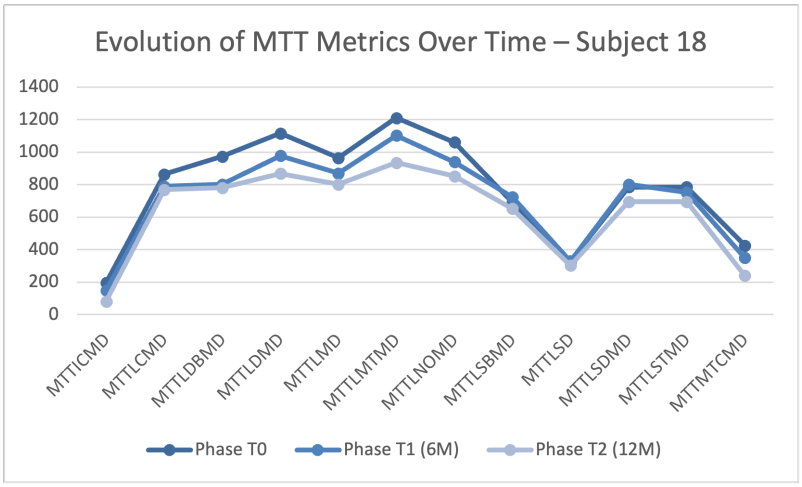


Figure 4.25: Evaluation of the variables MTTICMD, MTTLCMD, MTTLDBMD, MTTLDMD, MTTLMTMD, MTTLNOMD, MTTLSBMD, MTTLSD, MTTLSDMD, MTTLSTMD and MTTMTCMD during MTT test for the patient number 18P

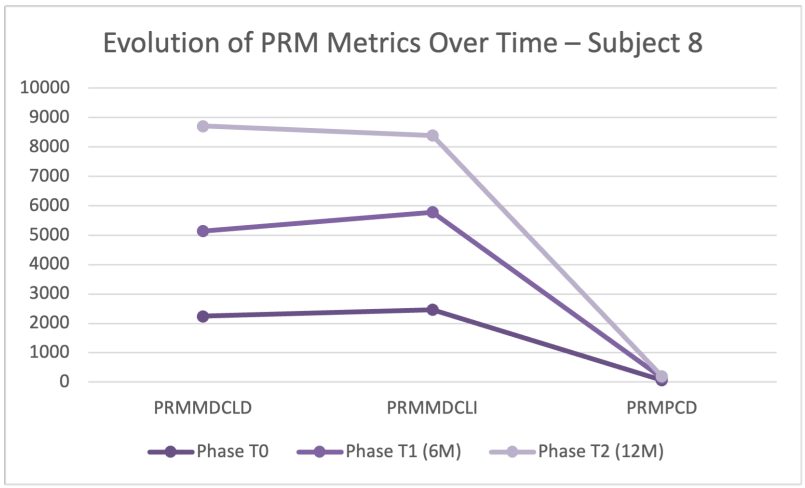


Figure 4.26: Evaluation of the variables PRMMDCLD, PRMMDCLI and PRMPCD during PRM test for the patient number 8P

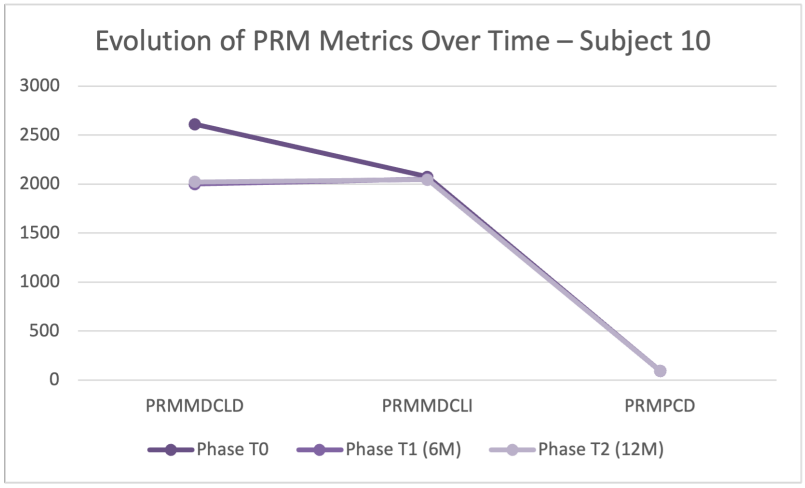


Figure 4.27: Evaluation of the variables PRMMDCLD, PRMMDCLI and PRMPCD during PRM test for the patient number 10P

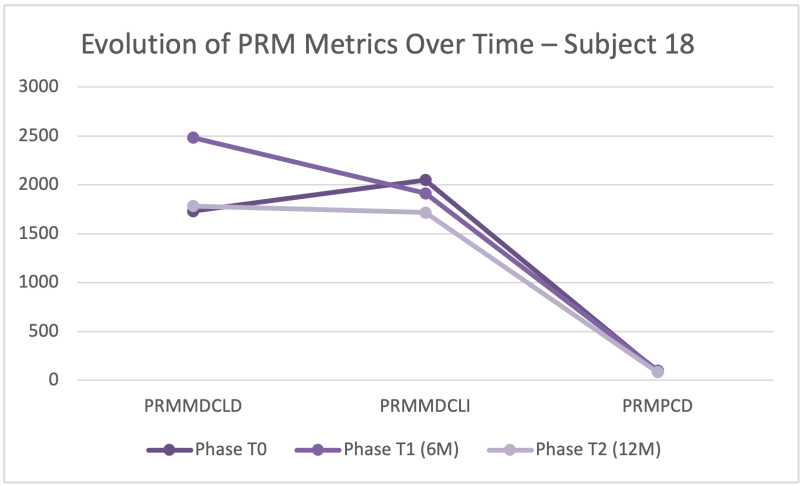


Figure 4.28: Evaluation of the variables PRMMDCLD, PRMMDCLI and PRMPCD during PRM test for the patient number 18P

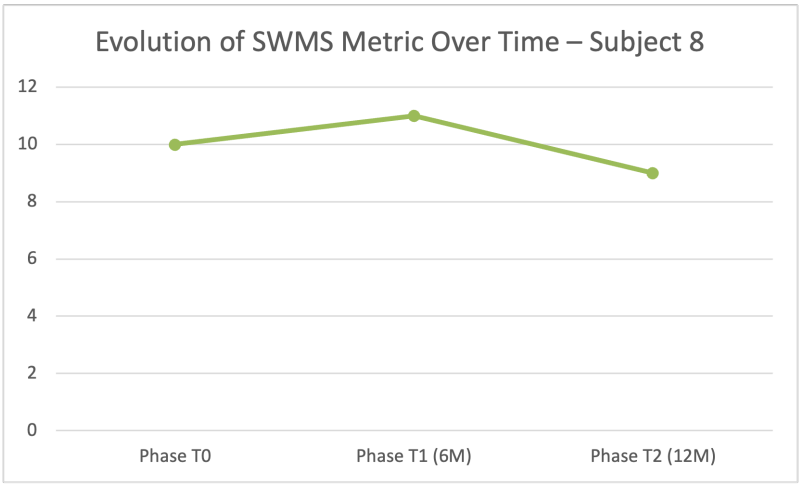


Figure 4.29: Evaluation of the variables SWMS during SWM test for the patient number 8P

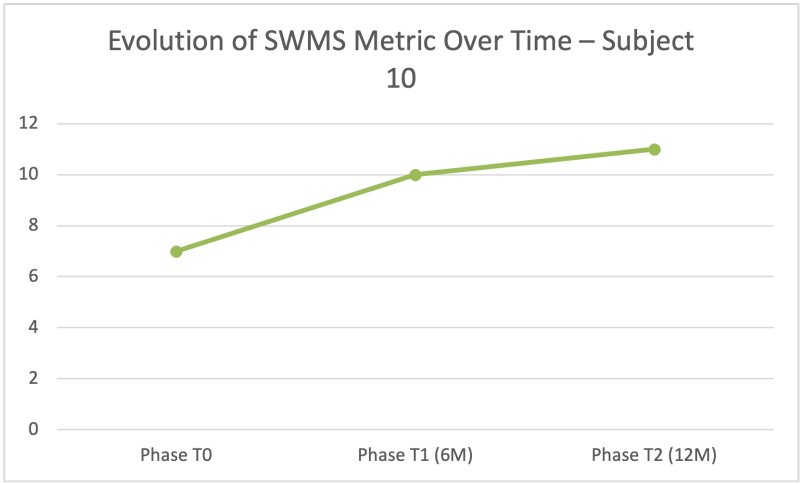


Figure 4.30: Evaluation of the variables SWMS during SWM test for the patient number 10P

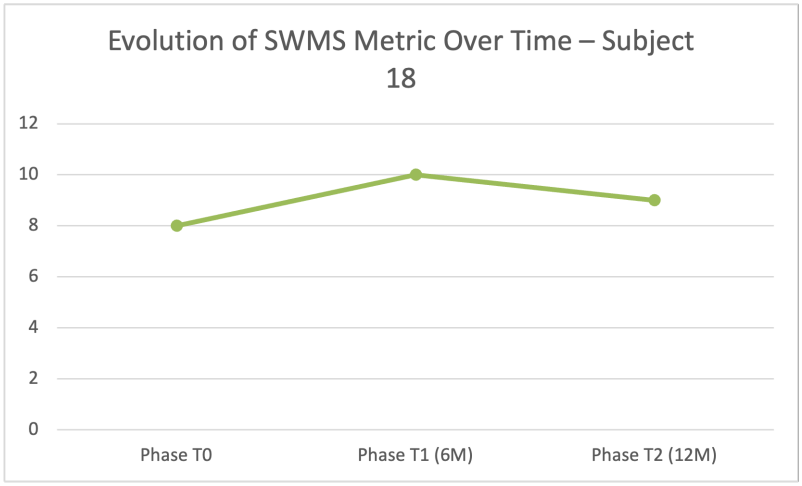


Figure 4.31: Evaluation of the variables SWMS during SWM test for the patient number 18P

Figures 4.20, 4.21 and 4.22 show plots for each subject in the aMCI group during the performance of the ERT test in the three experimental phases. In the subject identified with number 8, the average latency of the subject in selecting the correct emotion during test execution (ERTCRT) and the average latency of the subject in selecting a response (regardless of whether it is correct) after stimulus administration (ERTRT) are slightly increasing over the experimental phases. The total number of correct responses (ERTTH) remains almost unchanged. Figures 4.23, 4.24 and 4.25 show the trends during the performance of the MTT in the three time phases interspersed by a 6-month period. Here, a slight deterioration in performance is again evident in subject 8 as the values of incongrunet cost (MTTICMD), mean latency and standard deviations increase during T1 and T2. Even when performing the PRM test (figure 4.26), patient 8 showed a noticeable worsening in the variables of mean response latency in conditions in which the subject responds correctly when forced to a choice immediately (PRMMDCLI) and mean response latency when the subject responds correctly after being induced late to a choice (PRMMDCLD). The percentage of correct responses (PRMPC) remains stable. Patients 10 and 18 have a more variable trend, with some phases of improvement and other phases of worsening (e.g., in performing ERT, subject 10 manifests an increase in mean latency values in the experimental T1 phase, and then achieves a decreasing trend in the T2 phase). This behavior is common in all tests performed (ERT, PRM and MTT). In the SWM test, the SWMS parameter was

evaluated, which indicates the number of times the subject always starts a search pattern from the same box. In this case subjects 8 and 18 (4.29 and 4.31 take a variable trend, in which there is evidence of worsening in the first experimental phase and slight improvement in the second phase. While subject 10 (4.30 has a decreasing trend in performance over the three experimental phases.

4.2 Eye-tracker signal outcomes

The results obtained from statistical analysis using the Wilcoxon-Mann-Whitney test, described in the previous chapter, are shown in the following tables. The tables show the trends of the medians of the distributions of variables. Both show, for each variable, the mean *pm* standard deviation, the *p-value*, the *rank statistic* and the *h-value*. The table below assesses whether there is a significant difference between the two types of stimuli when performing the two phases:

Table 4.3: p-values, sign rank test statistic value and h-values obtained from Wilcoxon-Mann-Whitney for paired sample test application, considering comparisons among different tasks (Encoding vs Recall). Significance level of 5% is adopted to state if any variables distribution shows differences in their median value. The test statistic is defined as the smaller value between the sum of positive ranks and the sum of the negative ranks of the series of numbers considered.

Variable	Category	p-value	Statistic	Significant
numFix_median	Encoding	0.04009352	269.5	1
numFix_median	Recall	0.17363436	219.5	0
meanFixDur_median	Encoding	0.28525952	160.0	0
meanFixDur_median	Recall	0.02257281	119.5	1
stdFixDur_median	Encoding	0.04356356	273.0	1
stdFixDur_median	Recall	0.84101273	208.0	0
latencyFirstFix_median	Encoding	0.32406845	171.0	0
latencyFirstFix_median	Recall	0.57335894	190.0	0

Variable	Category	p-value	Statistic	Significant
numSac_median	Encoding	0.92455862	196.0	0
numSac_median	Recall	0.16258685	220.0	0
meanSacDist_median	Encoding	0.93529216	196.5	0
meanSacDist_median	Recall	0.16258685	220.0	0
maxDistSac_median	Encoding	0.03501066	278.0	1
maxDistSac_median	Recall	0.16258685	220.0	0
MeanAngVel_median	Encoding	0.28876706	160.0	0
MeanAngVel_median	Recall	0.02830843	281.0	1
visualEntropy_median	Encoding	0.21291211	244.0	0
visualEntropy_median	Recall	0.07255961	264.0	0

The table below shows whether there is a significant difference between the two phases (Encoding and Recall) during the administration of each type of stimulus:

Table 4.4: p-values, sign rank test statistic value and h-values obtained from Wilcoxon-Mann-Whitney for paired sample test application, considering comparisons among different type of stimulus (Image Personal vs Image Neffie). Significance level of 5% is adopted to state if any variables distribution shows differences in their median value. The test statistic is defined as the smaller value between the sum of positive ranks and the sum of the negative ranks of the series of numbers considered.

Variable	Category	p-value	Statistic	Significant
numFix_median	Image Neffie	1.011127e-08	400.0	1
numFix_median	Image Personal	2.817930e-08	391.5	1
meanFixDur_median	Image Neffie	5.730972e-05	348.0	1
meanFixDur_median	Image Personal	9.248054e-07	381.0	1
stdFixDur_median	Image Neffie	3.317146e-05	47.0	1
stdFixDur_median	Image Personal	2.587190e-06	26.0	1

Variable	Category	p-value	Statistic	Significant
latencyFirstFix_median	Image Neffie	7.742539e-02	248.0	0
latencyFirstFix_median	Image Personal	1.716024e-01	229.5	0
numSac_median	Image Neffie	2.991982e-08	390.0	1
numSac_median	Image Personal	2.940104e-05	345.0	1
meanSacDist_median	Image Neffie	2.991982e-08	390.0	1
meanSacDist_median	Image Personal	1.503400e-06	365.0	1
maxDistSac_median	Image Neffie	8.006545e-09	400.0	1
maxDistSac_median	Image Personal	1.514940e-08	400.0	1
MeanAngVel_median	Image Neffie	6.529000e-10	7.0	1
MeanAngVel_median	Image Personal	1.450889e-11	0.0	1
visualEntropy_median	Image Neffie	4.065980e-01	171.5	0
visualEntropy_median	Image Personal	9.944416e-02	140.0	0

The null hypothesis adopted is that the difference between the median of a variable measured in the Encoding test during exposure to the Neffie stimulus and the same value belonging to the Recall test during the same stimuli is zero. The tables indicate how many and which varaibles showed significant differences in their median values.

Next, the variables of the control and aMCI subjects were visualized using descriptive analysis techniques. Since the sample of patients with aMCI is small (n=3), this analysis can be considered exploratory, aimed at observing general trends without drawing statistically definitive conclusions. The following figures show the distributions obtained, where subjects 1, 2 and 3 (whose values are indicated by blue scatter points) indicate patients identified with 8P, 10P and 18P, respectively, and the orange line indicates the mean value of healthy subjects for the variable analyzed:

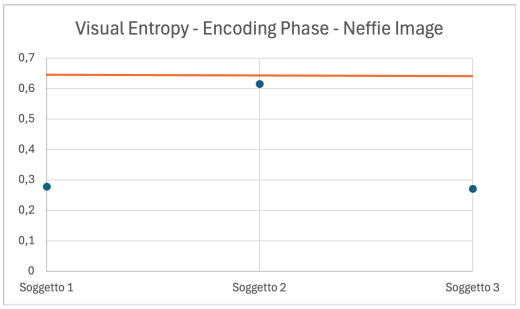


Figure 4.32: Distribution of individual values for the variable Visual Entropy during the Encoding Phase while viewing Neffie images, with one data point per patient. The solid line represents the average trend of the control group.

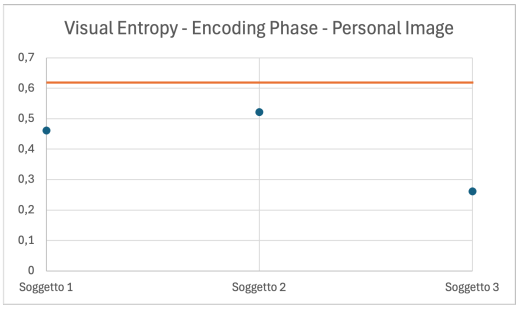


Figure 4.33: Distribution of individual values for the variable Visual Entropy during the Encoding Phase while viewing Personal images, with one data point per patient. The solid line represents the average trend of the control group.

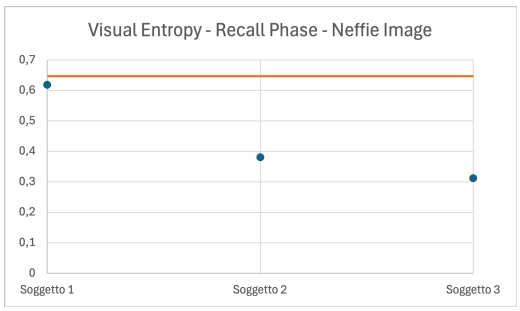


Figure 4.34: Distribution of individual values for the variable Visual Entropy during the Recall Phase while viewing Neffie images, with one data point per patient. The solid line represents the average trend of the control group.

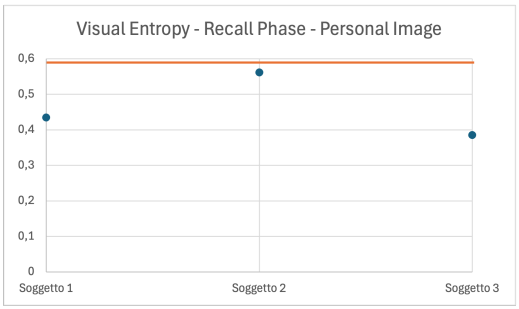


Figure 4.35: Distribution of individual values for the variable Visual Entropy during the Recall Phase while viewing Personal images, with one data point per patient. The solid line represents the average trend of the control group.

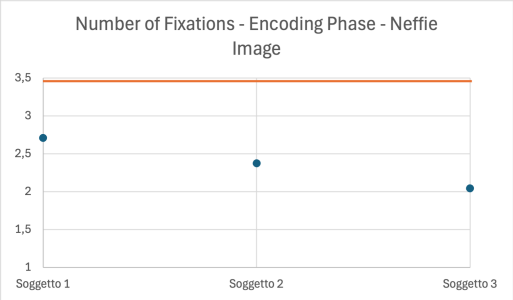


Figure 4.36: Distribution of individual values for the variable Number of Fixations during the Encoding Phase while viewing Neffie images, with one data point per patient. The solid line represents the average trend of the control group.

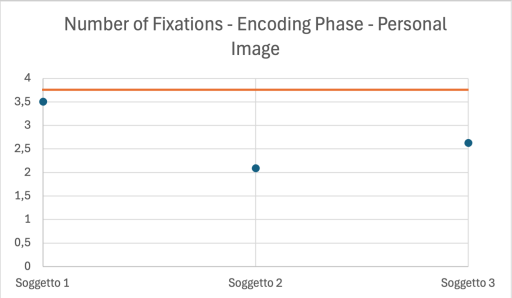


Figure 4.37: Distribution of individual values for the variable Number of Fixations during the Encoding Phase while viewing Personal images, with one data point per patient. The solid line represents the average trend of the control group.

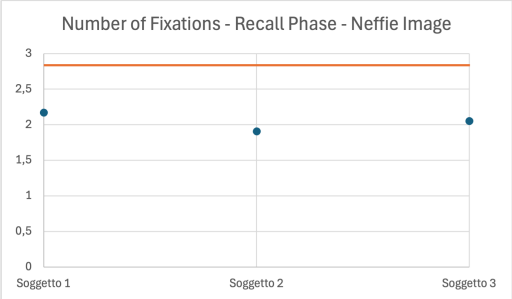


Figure 4.38: Distribution of individual values for the variable Number of Fixations during the Recall Phase while viewing Neffie images, with one data point per patient. The solid line represents the average trend of the control group.

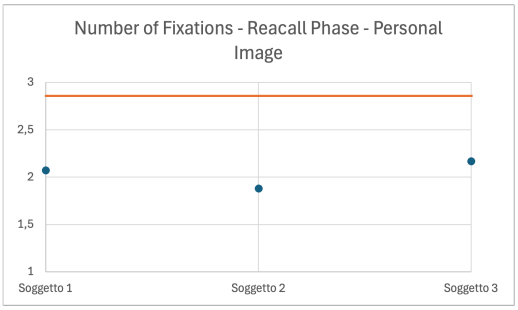


Figure 4.39: Distribution of individual values for the variable Number of Fixations during the Recall Phase while viewing Personal images, with one data point per patient. The solid line represents the average trend of the control group.

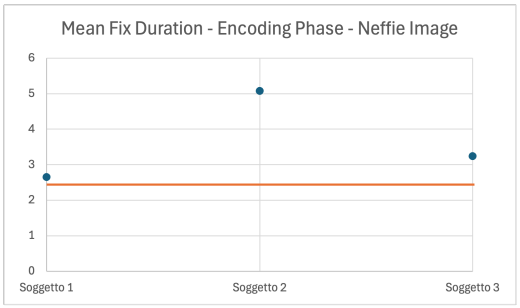


Figure 4.40: Distribution of individual values for the variable Mean duration of fixation during the Encoding Phase while viewing Neffie images, with one data point per patient. The solid line represents the average trend of the control group.

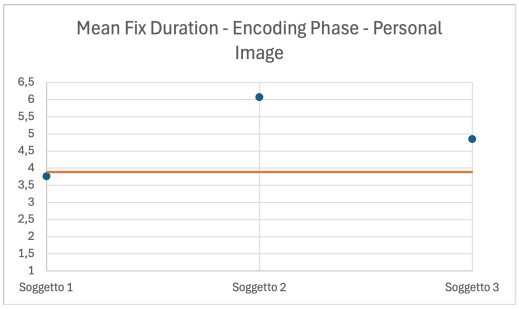


Figure 4.41: Distribution of individual values for the variable Mean duration of fixation during the Encoding Phase while viewing Personal images, with one data point per patient. The solid line represents the average trend of the control group.

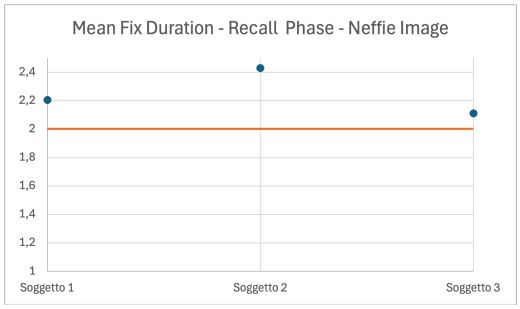


Figure 4.42: Distribution of individual values for the variable Mean duration of fixation during the Recall Phase while viewing Neffie images, with one data point per patient. The solid line represents the average trend of the control group.

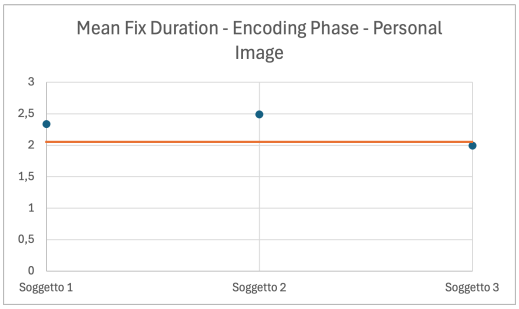


Figure 4.43: Distribution of individual values for the variable Mean duration of fixation during the Recall Phase while viewing Personal images, with one data point per patient. The solid line represents the average trend of the control group.

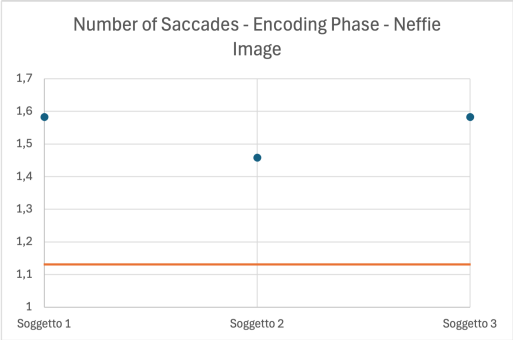


Figure 4.44: Distribution of individual values for the variable Number of saccades during the Encoding Phase while viewing Neffie images, with one data point per patient. The solid line represents the average trend of the control group.

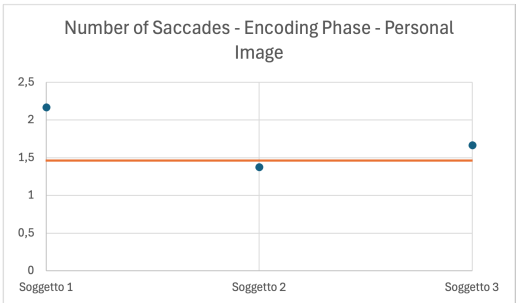


Figure 4.45: Distribution of individual values for the variable Number of saccades during the Encoding Phase while viewing Personal images, with one data point per patient. The solid line represents the average trend of the control group.

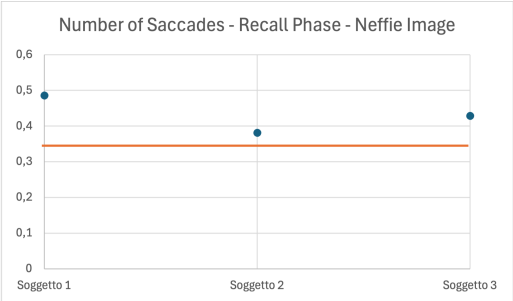


Figure 4.46: Distribution of individual values for the variable Number of saccades during the Recall Phase while viewing Neffie images, with one data point per patient. The solid line represents the average trend of the control group.

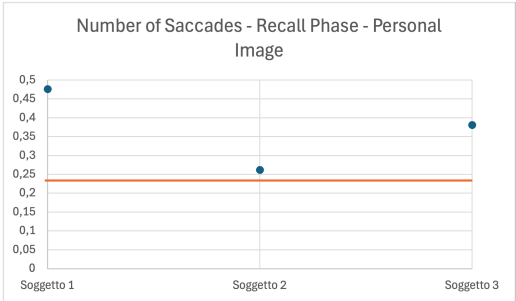


Figure 4.47: Distribution of individual values for the variable Number of saccades during the Recall Phase while viewing Personal images, with one data point per patient. The solid line represents the average trend of the control group.

These trends are obtained by evaluating the paramenters of individual subjects plotted together with a line representing the mean for that value in control subjects. The graphs are differentiated by task (Encoding or Recall) and by stimulus type (personal image or Neffie image).

Chapter 5

Discussion

The objective of this thesis project is to validate the use of the eye-tracker, integrated into a platform with several sensors, as a tool to support early detection of amnestic cognitive impairment, a condition that precedes Alzheimer's disease proper. The analysis of eye movement parameters was performed in two ways: - first, as described in the previous chapter, only control subjects (n=22) were evaluated using the Wilcoxon-Mann-Whitney statistical test. - In addition, given the difficulty in recruiting aMCI patients (n=3), an exploratory analysis aimed at visualizing general trends without statistical significance was performed. Statistical evaluation in the control group allowed a comparison between the different types of stimuli (Neffie images and personal images) and the different tasks performed (Encoding and Recall). In the case of evaluating the significant difference between the two phases (Encoding and Recall), the variable of the median number of fixations showed statistical significance that can be justified by a higher attentional level required by the Encoding phase. Higher visual attention had also been assessed in the Johansson, (2012)[76] study. For the Recall phase, a variable that has shown relevant statistical significance is the median of the mean duration of fixations. In fact, the median of the mean duration of fixations is higher than in the encoding phase as it increases the cognitive load related to the recall of stored images as analyzed in the study Damiano, C.(2020)[77]. Analysis of the significant differences between the two types of stimuli (personal images and Neffie images) during the two phases produced several significant variables highlighted in yellow in Table 4.4. According to the study conducted by Grill-Spector et al., 2004 and Summerfield et al., 2008, stimulus familiarity can influence attention and visual memory. These results were obtained exclusively in controls, but could serve as starting points for evaluating aMCI patients' visual exploration of personal stimuli. According to the study by Della Sala et al., 2012, patients with aMCI have a reduced ability to detect familiarity in stimuli. These results are consistent with neuropsychology studies: aMCI patients have a higher prevalence of depressive symptoms than healthy subjects. Therefore, it would be interesting to compare these results with those obtained with neuropsychological tests such as the Emotion Recognition Task (ERT) from the Cantab battery to assess whether the difficulties experienced by the patient in recognizing emotional expressions may be reflected in differences in oculomotor parameters. The second part of the eye-tracker signal analysis is focused on preliminary observation by focusing on the distribution of variables for individual aMCI patients compared with the mean of control subjects for the parameter under consideration. The variables evaluated in this case are:

- number of fixations
- mean duration of fixations
- number of saccades
- visual entropy

These variables were assessed by distinguishing between tasks in the experimental phase and the type of stimulus they are subjected to. Compared with the average trend of healthy subjects, patients with aMCI showed fewer stable fixations and shorter fixation durations. This trend is in line with what is defined in the literature by the study of Itti & Koch, 2001 and Matuso et al., 2013), according to which healthy subjects have a greater ability to focus attention on the proposed stimuli. This ability also results in a longer duration in fixations in control subjects who are able to maintain attention on a stimulus for a longer time than subjects with cognitive impairment. The aMCI patients individually showed fewer fixations and shorter duration of fixations. This behavior has been noted in the literature in the study by (Johansson et al., 2012) and explains, also, the presence of a greater number of saccades in patients. Saccades are essential for the focusing movement of attention while viewing an image and for visual selection. In aMCI patients this process might be altered, as analyzed in the study by Matuso et al. (2013). Finally, in Figures 4.32, 4.33, 4.34 and 4.35 , a reduced entropy is observed compared to the average of control subjects. Entropy represents variability in eye movements and in the distribution of fixations. Reduced entropy inidicates less spatial exploration. In order to validate these trends, an increase in sample number is needed so that a more robust statistical analysis can be performed. The results obtained offer interesting insights into visual exploration in patients with aMCI but have some limitations including possible varaible confounders such as educational level, residual ability and medication use that could influence the result. In addition to the analysis of ocular parameters, participants underwent the digitized battery of CANTAB cognitive tests for more detailed neuropsychological assessment. Longitudinal analysis performed on the control subjects showed no significant changes in any of the observed variables during the observation period, as indicated by the *p-value* greater than the significance threshold in the table 4.2. This suggests that, in the sample of healthy subjects, the cognitive functions examined (multitasking, spatial memory, emotion recognition, and visual memory) are stable in the short term. While no significant variations can be observed, it is important to

note that some variables over time take on higher performing values. For example, the trend in the percentage of correct answers in the PRM test (PRMPCD, Figure 4.15) increases over the selected time period indicating better performance. In contrast, the indicative variables of the average latency during test execution have a stable trend (MTTLCMD, Figure 4.2; MTTLDMD, Figure 4.4; MTTLMD, Figure 4.5; MTTLNOMD, Figure 4. 7) or decrease showing improvement (MTTICMD, Figure 4.1; MTTLMTMD, Figure 4.6; MTTLSD, Figure 4.9; PRMMDCLD, Figure 4.13; SWMS, Figure 4.16; ERTTH, Figure 4.19). This positive trend may be due to the learning effect described in the study by Karpicke & Roediger (2008)[78]. This effect occurs when subjects undergo multiple test sessions with the same test setting and thus over time, especially in a control group, could improve their performance. This improvement was not due to a change in cognitive function but simply to familiarity with the task performed. On the other hand, with regard to the single-subject analysis of aMCI subjects, significance was not displayed in a statistical sense, but it was possible to interpret the data according to the graphical display Patient 8 shows a slight steady deterioration in tests indicative of a progressive decline typical of aMCI. From the literature[79] it is known that aMCI is a developmental condition that leads to a gradual deterioration of cognitive function. Subject 8, in the ERT test shows an increase in response latency but no obvious change in the number of correct responses. This indicates that the subject is still able to recognize emotions, but due to cognitive slowing may take longer. Also in the PRM test, increased latency evidences greater difficulty in short-term memory retrieval. Executive functions assessed in the MTT test, moreover, show impairment over time in the case of subject 8 as an increase in the variables of inconsistency and mean latency is evident. Patients 10 and 18, on the other hand, show a more variable trend with phases of improvement and phases of worsening. This variability confirms the complexity of the aMCI condition, which presents instability in symptomatology and cognitive performance[79]. Future studies could include a larger sample number and longer follow-up period to maximize the potential of the Cantab test in monitoring cognitive function and help clinicians in psychological assessment aimed at early detection of mild cognitive disorder that could progress to AD.

Chapter 6

Conclusion

6.1 Summary of Contributions

The project presented in this thesis work represents a significant contribution in identifying and developing the potential of the NAP4A platform in the application of diagnostic biomarker of Alzheimer's disease in the preclinical stage, in individuals with Mild Cognitive Impairment (aMCI). The innovative NAP4A platform, developed by the research team at the Center for Advanced Technology in Health and Wellbeing, has been proposed as a tool to support the clinical assessment of biological signals of patients in the early stage of AD. The platform leverages the integration between different biological signals such as PPG, EEG, GSR, Eye-tracker and Webcam and the digitized neuropsychological test battery (CANTAB). The innovative key of the platform is the possibility of the simultaneous acquisition and processing of biomedical signals and the execution of cognitive tests. This makes it possible to overcome the limitations of traditional testing, facilitating early, affective and accurate diagnosis. The accuracy and robustness of the NAP4A platform proved to be remarkable during the acquisitions allowing the collection of a large amount of data, from different sensors, in real time. The study prior to mine, carried out in 2021, continued the photoplethysmographic signal analysis conducted on a control group by validating PPG signal processing. The signal I analyzed was that of the eye-tracker: following the acquisition and processing of the signal, longitudinal analysis performed on the control group and the patient group allowed us to compare the values with those found in the literature, to lay the foundation for future developments in early Alzheimer's diagnosis. Significant variables were identified that could be used as a discriminator between subjects with aMCI and healthy subjects.

6.2 Further Investigations and Improvements

The NAP4A platform and the approach developed in this thesis project have shown great potential for diagnosis in the preclinical stage of AD. However, the effectiveness and versatility of the platform can be increased by implementation of improvements. One of the major developments concerns the analysis of the electroencephalographic signal, which in the course of the study showed multiple issues related to signal transmission. These malfunctions caused a reduction in the robustness of the system in EEG signal

acquisition and integration as a large amount of data was lost. To reduce the inconveniences related to technical issues, it would be optimal to implement the platform by means of a more robust data transmission system, so as to allow a complete assessment of brain activities in MCI subjects and healthy ones. Further objective, made explicit in the protocol, concerns the application of artificial intelligence algorithms to improve the identification of meaningful patterns in the collected data. This would allow a reduction of time in signal post-processing and improved accuracy. Another goal concerns the expansion of the recruited sample. The protocol calls for enrolling approximately 30 control subjects and 20 aMCI patients. However, the current limited availability of aMCI subjects does not allow for statistical comparisons between the two groups and, consequently, for quantitative and meaningful results. In addition, longitudinal analysis of the data extracted from the two groups would be crucial for monitoring neurophysiological and cognitive characteristics over time. In conclusion, the results obtained in this study represent a huge step toward the realization of an innovative system for the early diagnosis of Alzheimer's disease that aims at an improvement in the quality of life of the patient and family members. Future developments related to platform optimization and databse expansion will help make the potential protocolized diagnostic system more accurate. Confirming its originality and applicability, the NAP4A2020 project received a patent in 2024.

Bibliography

[1] World Health Organization. *Dementia*. Accessed: 2025-03-16. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/dementia (cit. on pp. 3, 5).

[2] Osservatorio Demenze dell'Istituto Superiore di Sanità. *Definizione di Demenza e Disturbo Neurocognitivo*. Accessed: 2025-03-06. n.d. URL: https://www.demenze.it/it-schede-34-demenza_definizione# (cit. on p. 3).

[3] Michela Perrone. "Analisi quantitativa del 24S-idrossicolesterolo come applicazione clinica della LC-MS/MS alle malattie neurologiche: Alzheimer e demenza vascolare". Ph.D. thesis. Università degli Studi di Roma Tor Vergata, 2025. URL: https://art.torvergata.it/retrieve/e291c0d3-6fb2-cddb-e053-3a05fe0aa144/tesi%20dottorato_Michela%20Perrone.pdf (cit. on p. 3).

[4] Leon Stefanovski et al. "Bridging scales in Alzheimer's disease: biological framework for brain simulation with the virtual brain". In: *Frontiers in Neuroinformatics* 15 (2021), p. 630172 (cit. on pp. 4, 6, 8, 13, 83).

[5] John L Robinson, Maria M Corrada, Gabor G Kovacs, Myrna Dominique, Carrie Caswell, Sharon X Xie, Virginia M-Y Lee, Claudia H Kawas, and John Q Trojanowski. "Non-Alzheimer's contributions to dementia and cognitive resilience in The 90+ Study". In: *Acta neuropathologica* 136 (2018), pp. 377–388 (cit. on p. 4).

[6] Istituto Superiore di Sanità. *Alzheimer: 55 milioni di casi di demenza nel mondo, intervenendo su 14 fattori di rischio si potrebbero prevenire o ritardare la metà dei casi*. Accessed: 2025-03-16. 2025. URL: https://www.iss.it/en/-/alzheimer-55-milioni-di-casi-di-demenza-nel-mondo-intervenendo-su-14-fattori-di-rischio-si-potrebbero-prevenire-o-ritardare-la-meta-dei-casi (cit. on p. 5).

[7] Eurostat. *Eurostat Statistical Books - Statistics Explained*. Accessed: 2025-03-07. 2020. URL: https://ec.europa.eu/eurostat/web/products-statistical-books/-/ks-02-20-655 (cit. on p. 5).

[8] Chiara Zucchella, Elena Sinforiani, Stefano Tamburin, Angela Federico, Elisa Mantovani, Sara Bernini, Roberto Casale, and Michelangelo Bartolo. "The multidisciplinary approach to Alzheimer's disease and dementia. A narrative review of non-pharmacological treatment". In: *Frontiers in neurology* 9 (2018), p. 1058 (cit. on p. 7).

[9] Alzheimer's Association. *Alzheimer's Association*. Accessed: 2025-03-07. 2025. URL: https://www.alz.org (cit. on pp. 7, 10).

[10] Rihui Li, Guoxing Rui, Wei Chen, Sheng Li, Paul E Schulz, and Yingchun Zhang. "Early detection of Alzheimer's disease using non-invasive near-infrared spectroscopy". In: *Frontiers in aging neuroscience* 10 (2018), p. 366 (cit. on p. 7).

[11] Peter R Mouton, Lee J Martin, Michael E Calhoun, Gloria Dal Forno, and Donald L Price. "Cognitive decline strongly correlates with cortical atrophy in Alzheimer's dementia". In: *Neurobiology of aging* 19.5 (1998), pp. 371–377 (cit. on pp. 8, 9).

[12] B. Dubois et al. *Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria*. Accessed: 2025-03-07. 2007. URL: https://www.bakardjian.com/work/publications_other/Dubois__Research%20criteria%20for%20the%20diagnosis%20of%20AD%20-%20rev%20the%20NINCDS-ADRDA%20crit%20(2007).pdf (cit. on pp. 11, 12).

[13] Ove Almkvist. "Neuropsychological features of early Alzheimer's disease: preclinical and clinical stages". In: *Acta Neurologica Scandinavica* 94.S165 (1996), pp. 63–71 (cit. on pp. 13, 14).

[14] Willem S Eikelboom et al. "Neuropsychiatric and cognitive symptoms across the Alzheimer disease clinical spectrum: cross-sectional and longitudinal associations". In: *Neurology* 97.13 (2021), e1276–e1287 (cit. on p. 13).

[15] Author Name. *Title of the Document*. Accessed: 2025-03-07. 2025. URL: https://research-repository.griffith.edu.au/server/api/core/bitstreams/47440b20-4029-45d2-8175-80454538b6a9/content (cit. on p. 13).

[16] Xuqian Li and Junjing Wang. “Abnormal neural activities in adults and youths with major depressive disorder during emotional processing: a meta-analysis”. In: *Brain imaging and behavior* 15 (2021), pp. 1134–1154 (cit. on p. 14).

[17] Suzanne Penna. “Cognitive and emotional dysfunction in mild cognitive impairment”. In: *Clinics in geriatric medicine* 29.4 (2013), pp. 773–789 (cit. on p. 15).

[18] Lívia Spíndola and Sonia Maria Dozzi Brucki. “Prospective memory in Alzheimer’s disease and Mild Cognitive Impairment”. In: *Dementia & Neuropsychologia* 5.2 (2011), pp. 64–68 (cit. on p. 15).

[19] Noll L Campbell, Fred Unverzagt, Michael A LaMantia, Babar A Khan, and Malaz A Boustani. “Risk factors for the progression of mild cognitive impairment to dementia”. In: *Clinics in geriatric medicine* 29.4 (2013), p. 873 (cit. on p. 16).

[20] Frank Jessen et al. “A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer’s disease”. In: *Alzheimer’s & dementia* 10.6 (2014), pp. 844–852 (cit. on p. 16).

[21] Ronald C Petersen. “Early diagnosis of Alzheimer’s disease: is MCI too late?” In: *Current Alzheimer Research* 6.4 (2009), pp. 324–330 (cit. on p. 16).

[22] Jason Weller and Andrew Budson. “Current understanding of Alzheimer’s disease diagnosis and treatment”. In: *F1000Research* 7 (2018), F1000–Faculty (cit. on p. 17).

[23] John C Morris, Kaj Blennow, Lutz Frölich, Agneta Nordberg, Hilkka Soininen, Gunhild Waldemar, L-O Wahlund, and B Dubois. “Harmonized diagnostic criteria for Alzheimer’s disease: recommendations”. In: *Journal of internal medicine* 275.3 (2014), pp. 204–213 (cit. on p. 17).

[24] Lei Zhuang, Yan Yang, and Jianqun Gao. “Cognitive assessment tools for mild cognitive impairment screening”. In: *Journal of neurology* 268.5 (2021), pp. 1615–1622 (cit. on p. 19).

[25] Ronald C Petersen, Barbara Caracciolo, Carol Brayne, Serge Gauthier, Vesna Jelic, and Laura Fratiglioni. “Mild cognitive impairment: a concept in evolution”. In: *Journal of internal medicine* 275.3 (2014), pp. 214–228 (cit. on pp. 19–21).

[26] Alisoun Milne, Alison Culverwell, Reinhard Guss, Jackie Tuppen, and R Whelton. “Screening for dementia in primary care: a review of the use, efficacy and quality of measures”. In: *International psychogeriatrics* 20.5 (2008), pp. 911–926 (cit. on p. 20).

[27] Javier Arbizu et al. “Clinical utility of FDG-PET for the clinical diagnosis in MCI”. In: *European journal of nuclear medicine and molecular imaging* 45 (2018), pp. 1497–1508 (cit. on p. 20).

[28] Ronald C Petersen. “Mild cognitive impairment as a diagnostic entity”. In: *Journal of internal medicine* 256.3 (2004), pp. 183–194 (cit. on p. 21).

[29] María del Carmen Díaz-Mardomingo, Sara García-Herranz, Raquel Rodríguez-Fernández, César Venero, and Herminia Peraita. “Problems in classifying mild cognitive impairment (MCI): one or multiple syndromes?” In: *Brain sciences* 7.9 (2017), p. 111 (cit. on p. 22).

[30] Jessica Beltrán, Mireya S García-Vázquez, Jenny Benois-Pineau, Luis Miguel Gutierrez-Robledo, and Jean-François Dartigues. “Computational techniques for eye movements analysis towards supporting early diagnosis of Alzheimer’s disease: a review”. In: *Computational and mathematical methods in medicine* 2018.1 (2018), p. 2676409 (cit. on p. 23).

[31] Adam L Boxer, Siobhan Garbutt, Katherine P Rankin, Joanna Hellmuth, John Neuhaus, Bruce L Miller, and Stephen G Lisberger. “Medial versus lateral frontal lobe contributions to voluntary saccade control as revealed by the study of patients with frontal lobe degeneration”. In: *Journal of Neuroscience* 26.23 (2006), pp. 6354–6363 (cit. on p. 23).

[32] Nicolas Noiret, Nicolas Carvalho, Éric Laurent, Gilles Chopard, Mickaël Binetruy, Magali Nicolier, Julie Monnin, Eloi Magnin, and Pierre Vandel. “Saccadic eye movements and attentional control in Alzheimer’s disease”. In: *Archives of Clinical Neuropsychology* 33.1 (2018), pp. 1–13 (cit. on pp. 23, 24).

[33] Samuel B Hutton. “Cognitive control of saccadic eye movements”. In: *Brain and cognition* 68.3 (2008), pp. 327–340 (cit. on p. 23).

[34] Aristea Ladas, Christos Frantzidis, Panagiotis Bamidis, and Ana B Vivas. “Eye blink rate as a biological marker of mild cognitive impairment”. In: *International Journal of Psychophysiology* 93.1 (2014), pp. 12–16 (cit. on p. 24).

[35] Sidney Katz, Amasa B Ford, Roland W Moskowitz, Beverly A Jackson, and Marjorie W Jaffe. “Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychosocial function”. In: *jama* 185.12 (1963), pp. 914–919 (cit. on p. 34).

[36] Neuropsicologia Web. *Valutazione funzionale della demenza: Scale di autonomia di vita quotidiana (Scala ADL, IADL, AAIDL)*. Accesso: 6 Marzo 2025. 2025. URL: https://www.neuropsicologiaweb.it/index.php/neuropsicologia/valutazione/85-valutazione-funzionale-della-demenza-scale-di-autonomia-di-vita-quotidiana-scala-adl-iadl-aadl?start=1&utm (cit. on pp. 34–36).

[37] SAM Sikkes, ESM De Lange-de Klerk, YAL Pijnenburg, and P Scheltens. "A systematic review of Instrumental Activities of Daily Living scales in dementia: room for improvement". In: *Journal of Neurology, Neurosurgery & Psychiatry* 80.1 (2009), pp. 7–12 (cit. on p. 34).

[38] Vladimir Hachinski, Shahram Oveisgharan, A Kimball Romney, and William R Shankle. "Optimizing the Hachinski ischemic scale". In: *Archives of neurology* 69.2 (2012), pp. 169–175 (cit. on p. 34).

[39] Akhilesh Vyas, Fotis Aisopos, Maria-Esther Vidal, Peter Garrard, and George Paliouras. "Calibrating mini-mental state examination scores to predict misdiagnosed dementia patients". In: *Applied Sciences* 11.17 (2021), p. 8055 (cit. on p. 34).

[40] Kimberly R Chapman, Hanaan Bing-Canar, Michael L Alosco, Eric G Steinberg, Brett Martin, Christine Chaisson, Neil Kowall, Yorghos Tripodis, and Robert A Stern. "Mini Mental State Examination and Logical Memory scores for entry into Alzheimer's disease trials". In: *Alzheimer's research & therapy* 8 (2016), pp. 1–11 (cit. on p. 35).

[41] Eli Vakil and Haya Blachstein. "Rey auditory-verbal learning test: structure analysis". In: *Journal of clinical psychology* 49.6 (1993), pp. 883–890 (cit. on p. 35).

[42] Center for Traumatic Brain Injury. *Trail Making Test (TMT) - Italian Version*. Accessed: 2025-03-06. 2025. URL: https://www.center-tbi.eu/files/approved-translations/Italian/ITALIAN_TMT.pdf (cit. on p. 36).

[43] Francesca Fiore. *Effetto Stroop - Il test di Stroop in Psicologia Sperimentale*. Accessed: 2025-03-06. 2018. URL: https://www.stateofmind.it/2018/06/effetto-stroop-psicologia/ (cit. on p. 36).

[44] Jane H Cerhan, Robert J Ivnik, Glenn E Smith, Eric C Tangalos, Ronald C Petersen, and Bradley F Boeve. "Diagnostic utility of letter fluency, category fluency, and fluency difference scores in Alzheimer's disease". In: *The Clinical Neuropsychologist* 16.1 (2002), pp. 35–42 (cit. on p. 36).

[45] Cambridge Cognition. *Cambridge Cognition*. Accessed: 2025-03-06. 2025. URL: https://cambridgecognition.com (cit. on pp. 36, 41).

[46] NeuroPSY. *BORB Test - Tavola 02*. Accessed: 2025-03-06. 2025. URL: https://www.neuropsy.it/test/borb/02.html (cit. on p. 36).

[47] Ordine degli Psicologi FVG. *WISC-IV - Contenuti*. Accessed: 2025-03-06. 2014. URL: https://www.ordinepsicologifvg.it/images/eventi/20141213_wiscIV/contenuti.pdf (cit. on p. 37).

[48] Alessandra Dodich et al. "A novel task assessing intention and emotion attribution: Italian standardization and normative data of the Story-based Empathy Task". In: *Neurological Sciences* 36 (2015), pp. 1907–1912 (cit. on p. 37).

[49] Janine Diehl-Schmid, Corina Pohl, Carolin Ruprecht, Stefan Wagenpfeil, Hans Foerstl, and Alexander Kurz. "The Ekman 60 Faces Test as a diagnostic instrument in frontotemporal dementia". In: *Archives of Clinical Neuropsychology* 22.4 (2007), pp. 459–464 (cit. on p. 37).

[50] Ignacio Montorio and Maria Izal. "The Geriatric Depression Scale: a review of its development and utility". In: *International psychogeriatrics* 8.1 (1996), pp. 103–112 (cit. on p. 37).

[51] Gordon Jackson-Koku. "Beck depression inventory". In: *Occupational medicine* 66.2 (2016), pp. 174–175 (cit. on p. 37).

[52] Andras N Zsido, Szidalisz A Teleki, Krisztina Csokasi, Sandor Rozsa, and Szabolcs A Bandi. "Development of the short version of the spielberger state—trait anxiety inventory". In: *Psychiatry research* 291 (2020), p. 113223 (cit. on p. 37).

[53] Donald J Connor, Marwan N Sabbagh, and Jeffery L Cummings. "Comment on administration and scoring of the Neuropsychiatric Inventory in clinical trials". In: *Alzheimer's & Dementia* 4.6 (2008), pp. 390–394 (cit. on p. 38).

[54] Michela Leocadi et al. "Dual-task gait training improves cognition and resting-state functional connectivity in Parkinson's disease with postural instability and gait disorders". In: *Journal of Neurology* 271.4 (2024), pp. 2031–2041 (cit. on p. 38).

[55] Monica Luciana and Charles A Nelson. "Assessment of neuropsychological function through use of the Cambridge Neuropsychological Testing Automated Battery: performance in 4-to 12-year-old children". In: *Developmental neuropsychology* 22.3 (2002), pp. 595–624 (cit. on p. 38).

[56] Katherine V Wild and Erica D Musser. “The Cambridge Neuropsychological Test Automated Battery in the assessment of executive functioning”. In: *Handbook of executive functioning*. Springer, 2013, pp. 171–190 (cit. on p. 39).

[57] KU Leuven - PASH Research Group. *CANTAB Connect Research Test Administration User Guide*. Accessed: 2025-03-06. 2025. URL: https://gbiomed.kuleuven.be/english/research/50000737/research/pash/pasoca/documenten/CANTAB-Connect-Research-Test-Administration-User-Guide.pdf (cit. on pp. 39, 102).

[58] Shimmer Sensing. *GSR User Guide*. Accessed: 2025-03-06. 2025. URL: https://shimmersensing.com/wp-content/docs/support/documentation/GSR_User_Guide_rev1.13.pdf (cit. on pp. 45, 47).

[59] Antonella Sirianni. *Testi*. Accessed: 2025-03-06. 2016. URL: https://etd.adm.unipi.it/theses/available/etd-01222016-100019/unrestricted/Testi_Antonella_Sirianni.pdf (cit. on p. 46).

[60] Toshiyo Tamura, Yuka Maeda, Masaki Sekine, and Masaki Yoshida. “Wearable photoplethysmographic sensors—past and present”. In: *Electronics* 3.2 (2014), pp. 282–302 (cit. on p. 48).

[61] Emotiv. *Emotiv Insight*. Accessed: 2025-03-06. 2025. URL: https://www.emotiv.com/products/insight?srsltid=AfmBOoorWgcIQ1CXQ6_LNZtEvEiI7xwtIob15fKuFVr0psCcKhOuvynY (cit. on p. 48).

[62] Tobii. *How Do Eye Trackers Work?* Accessed: 2025-03-06. n.d. URL: https://www.tobii.com/resource-center/learn-articles/how-do-eye-trackers-work (cit. on pp. 50, 51, 55).

[63] Roy S Hessels, Diederick C Niehorster, Marcus Nyström, Richard Andersson, and Ignace TC Hooge. “Is the eye-movement field confused about fixations and saccades? A survey among 124 researchers”. In: *Royal Society open science* 5.8 (2018), p. 180502 (cit. on p. 52).

[64] Rajgunesh. *Cubic Spline Interpolation*. Accessed: 2023-03-09. 2023. URL: https://www.rajgunesh.com/resources/downloads/numerical/cubicsplineinterpol.pdf (cit. on p. 67).

[65] Anuradha Kar and Peter Corcoran. “A review and analysis of eye-gaze estimation systems, algorithms and performance evaluation methods in consumer platforms”. In: *IEEE Access* 5 (2017), pp. 16495–16519 (cit. on pp. 69–72).

[66] Tobii Pro. *Calibration*. Accessed: 2023-03-09. 2023. URL: https://developer.tobiipro.com/commonconcepts/calibration.html (cit. on p. 69).

[67] Sileye O Ba and Jean-Marc Odobez. "Multiperson visual focus of attention from head pose and meeting contextual cues". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.1 (2010), pp. 101–116 (cit. on pp. 69, 70).

[68] Kirsten A Dalrymple, Marie D Manner, Katherine A Harmelink, Elayne P Teska, and Jed T Elison. "An examination of recording accuracy and precision from eye tracking data from toddlerhood to adulthood". In: *Frontiers in psychology* 9 (2018), p. 803 (cit. on p. 73).

[69] Marcus Nyström and Kenneth Holmqvist. "An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data". In: *Behavior research methods* 42.1 (2010), pp. 188–204 (cit. on pp. 75, 76, 79, 80).

[70] Mehedi H Raju, Lee Friedman, Troy M Bouman, and Oleg V Komogortsev. "Filtering eye-tracking data from an eyelink 1000: Comparing heuristic, savitzky-golay, iir and fir digital filters". In: *Journal of Eye Movement Research* 14.3 (2023), pp. 10–16910 (cit. on p. 76).

[71] Sarmad Al-gawwam and Mohammed Benaissa. "Robust eye blink detection based on eye landmarks and Savitzky–Golay filtering". In: *Information* 9.4 (2018), p. 93 (cit. on p. 76).

[72] Pieter Blignaut. "Fixation identification: The optimum threshold for a dispersion algorithm". In: *Attention, Perception, & Psychophysics* 71 (2009), pp. 881–895 (cit. on pp. 81, 83, 84).

[73] Dario D Salvucci and Joseph H Goldberg. "Identifying fixations and saccades in eye-tracking protocols". In: *Proceedings of the 2000 symposium on Eye tracking research & applications*. 2000, pp. 71–78 (cit. on pp. 83, 84).

[74] Samuel Sanford Shapiro and Martin B Wilk. "An analysis of variance test for normality (complete samples)". In: *Biometrika* 52.3-4 (1965), pp. 591–611 (cit. on p. 94).

[75] IBM. *Linear Mixed Models*. Accessed: 2025-03-10. 2025. URL: https://www.ibm.com/docs/it/spss-statistics/saas?topic=models-linear-mixed-model (cit. on p. 95).

[76] Roger Johansson, Jana Holsanova, Richard Dewhurst, and Kenneth Holmqvist. “Eye movements during scene recollection have a functional role, but they are not reinstatements of those produced during encoding.” In: *Journal of Experimental Psychology: Human perception and performance* 38.5 (2012), p. 1289 (cit. on p. 123).

[77] Rosa Angela Fabio, Iconia Gullà, Antonino Errante, et al. “Emotions and eye movements: Eye tracker and mnestic parameters”. In: *Memory consolidation* (2015), pp. 235–258 (cit. on p. 123).

[78] Henry L Roediger and Andrew C Butler. “The critical role of retrieval practice in long-term retention”. In: *Trends in cognitive sciences* 15.1 (2011), pp. 20–27 (cit. on p. 125).

[79] Eva Arnáiz and Ove Almkvist. “Neuropsychological features of mild cognitive impairment and preclinical Alzheimer’s disease”. In: *Acta Neurologica Scandinavica* 107 (2003), pp. 34–41 (cit. on p. 125).

[80] Emma Nichols et al. “Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019”. In: *The Lancet Public Health* 7.2 (2022), e105–e125.

[81] Author(s) Name. “Title of the Article”. In: *F1000Research* 7 (2025). Accessed: 2025-03-07, p. 1161. URL: `https://f1000research.com/articles/7-1161`.

[82] Smilja Stokanović, Vukašin Spasojević, Ilija Tanasković, Jaka Sodnik, and Nadica Miljković. “A Computational Method for Saccade Velocity”. In: *2024 IEEE 24th International Conference on Bioinformatics and Bioengineering (BIBE)*. IEEE. 2024, pp. 1–8.

Dedications

At the conclusion of this academic journey,I would like to express my sincerest thanks to all the people who supported and guided me in the realization of this master's thesis. First, I warmly thank my supervisor, Professor Filippo Molinari, for his constant availability. I also thank the team at San Raffaele Hospital, especially, co-advisors Ing. Alberto Sanna, Ing. Alessandra Gargiulo and Ing. Alice Leoporini for the resources provided and logistical support that greatly facilitated the realization of this thesis.