### TABLE OF CONTENTS

1. **Neuromolecular and Cognitive Background of Alzheimer’s Disease and Mild Cognitive Impairment, with Focus on the Role Played by the Apolipoprotein E Gene**  
   - General Overview and Objectives ........................................................................................................... 4  
   - 1.1 Alzheimer’s Disease: Theories and Clinical Relevance ........................................................................ 6  
     - 1.1.1 Main Hypotheses .......................................................................................................................... 6  
     - 1.1.2 Main AD Peptidic Biomarkers ......................................................................................................... 10  
   - 1.2 Multidimensional Progression of AD Pathology ................................................................................ 15  
   - 1.3 MCI as a Snapshot from Normality to Dementia ................................................................................. 18  
     - 1.3.1 MCI Definition ............................................................................................................................. 18  
     - 1.3.2 MCI and AD-Related Proteins ....................................................................................................... 20  
     - 1.3.3 MCI and Brain Structure ................................................................................................................ 23  
     - 1.3.4 MCI and Brain Function .................................................................................................................. 26  
     - 1.3.5 MCI and Cognitive Function .......................................................................................................... 32  
     - 1.3.6 MCI and Other Investigational Techniques .................................................................................... 35  
     - 1.3.7 MCI and the Timeline of Biomarkers ............................................................................................ 36  
   - 1.4 The Role of the Apolipoprotein E Genotype from Health to Pathology ........................................ 39  
     - 1.4.1 The APOE Gene ............................................................................................................................ 39  
     - 1.4.2 APOE Genotype and Imaging Genetics .......................................................................................... 42  

2. **Physical and Cognitive Stimulation as Interventions to Delay Normal and Abnormal Ageing Processes**  
   - 2.1 Physical Activity ............................................................................................................................... 62  
     - 2.1.1 Introduction ................................................................................................................................. 62  
     - 2.1.2 Physical Activity, Brain and Cognition in Observational Designs ................................................. 66  
     - 2.1.3 Physical Activity, Brain and Cognition in Exercise Interventions ................................................. 73  
     - 2.1.4 Conclusion ................................................................................................................................. 80  
   - 2.2 Cognitive Activity ............................................................................................................................. 82  
     - 2.2.1 Introduction ................................................................................................................................. 82  
     - 2.2.2 Highlights on Far Transfer in Healthy Adults ................................................................................. 84  
     - 2.2.3 Cognitive Activities, Brain and Cognition in Observational Designs ........................................... 87  
     - 2.2.4 Cognitive Activity and Cognition in Intervention Trials ............................................................... 89  
     - 2.2.5 Conclusion ................................................................................................................................. 98  

3. **Aims and Objectives** ......................................................................................................................... 99
4 SHORT-TERM TREADMILL EXERCISE, NEUROCOGNITIVE IMPACT, AND THE ROLE OF APOLIPOPROTEIN E AND STRESS RESPONSE ................................................................. 102

4.1 EXPERIMENT 1 - THE IMPACT OF ACUTE EXERCISE ON BRAIN AND COGNITION .... 103
   4.1.1 INTRODUCTION ....................................................................................................... 103
   4.1.2 METHODS ............................................................................................................... 109
   4.1.3 PROCEDURE .......................................................................................................... 112
   4.1.4 RESULTS ............................................................................................................... 114
   4.1.5 DISCUSSION ......................................................................................................... 118

4.2 EXPERIMENT 2 - THE ROLE OF STRESS RESPONSE .................................................... 120
   4.2.1 INTRODUCTION ....................................................................................................... 120
   4.2.2 METHODS ............................................................................................................... 124
   4.2.3 PROCEDURE .......................................................................................................... 125
   4.2.4 RESULTS ............................................................................................................... 125
   4.2.5 DISCUSSION ......................................................................................................... 128

4.3 EXPERIMENT 3 - THE ROLE OF THE APOLIPOPROTEIN E GENOTYPE ...................... 130
   4.3.1 INTRODUCTION ....................................................................................................... 130
   4.3.2 METHODS ............................................................................................................... 131
   4.3.3 PROCEDURE .......................................................................................................... 132
   4.3.4 RESULTS ............................................................................................................... 133
   4.3.5 DISCUSSION ......................................................................................................... 135

4.4 GENERAL DISCUSSION ................................................................................................. 137

5 INTENSIVE COGNITIVE STIMULATION AS NON-PHARMACOLOGICAL TREATMENT TO RESTORE REGIONAL CONNECTIVITY .................................................................. 141

5.1 A NETWORK-BASED PROGRAMME OF COGNITIVE TRAINING ON A SAMPLE OF AGEING ADULTS ........................................................................................................... 143
   5.1.1 EXPERIMENT 4 - EFFECTS ON BRAIN FUNCTION .............................................. 143
   5.1.2 EXPERIMENT 5 - EFFECTS ON BRAIN STRUCTURE ........................................... 161
   5.1.3 EXPERIMENT 6 - EFFECTS ON COGNITIVE FUNCTIONS ................................... 170

5.2 A NETWORK-BASED PROGRAMME OF COGNITIVE TRAINING ON A SAMPLE OF MCI 180
   5.2.1 EXPERIMENT 7 - EFFECT ON BRAIN FUNCTION .............................................. 180
   5.2.2 EXPERIMENT 8 - EFFECTS ON BRAIN STRUCTURE .......................................... 186
   5.2.3 EXPERIMENT 9 - EFFECT ON COGNITION ....................................................... 190

5.3 EXPERIMENT 10 - THE ROLE OF THE APOE GENE AS MEDIATOR OF BENEFIT ........ 196
   5.3.1 INTRODUCTION ....................................................................................................... 196
   5.3.2 METHODS ............................................................................................................... 197
LIST OF ABBREVIATIONS

18F-MPPF=4-[F-18] fluoro-N-[2-[4-(2-methoxyphenyl)-1-piperazinyl] ethyl]-N-[2-(pyridinil) benzamide (Page 13); 18F-FDDNP=18F-(2-(1-(2-[18F] fluorophenyl) (methyl) amino) 2-naphthyl) ethyldiene (Page 13); Aβ= Beta Amyloid (Page 6); ACC=Anterior Cingulate Cortex (Page 29); ACTIVE=Advanced Cognitive Training for Independent and Vital Elderly (Page 170); AD=Alzheimer’s Disease (Page 4); ADL=Activities of Daily Living (Page 18); ADP=Adenosine DiPhosphate (Page 64); AE=Acute Exercise (Page 63); ApoE=Apolipoprotein E (Page 5); APP=Amyloid Precursor Protein (Page 6); AS=Acute Stress (Page 121); ATP=Adenosine TriPhosphate (Page 64); BOLD=Blood-Oxygen-Level-Dependent (Page 29); CA=Cognitive Activity (Page 62); CA1=Cornu Ammonis 1 (Page 11); CDR=Clinical Dementia Rating (Page 19); CE=Chronic Exercise (Page 63); ChEI=Cholinesterase Inhibitor (Page 95); CR=Cognitive Reserve (Page 82); cRPE=Categorical Rate of Perceived Exertion (Page 114); CSF=Cerebrospinal Fluid (Page 7); DLPFC=Dorsolateral Prefrontal Cortex (Page 104); DMN=Default Mode Network (Page 30); DSMT=Digit Symbol Modalities Test (Page 33); DTI=Diffusion Tensor Imaging (Page 35); EEG=Electroencephalography (Page 35); FA=Fractional Anisotropy (Page 35); FAB=Frontal Assessment Battery (Page 78); FDG=Fluoro-D-Glucose (Page 26); fMRI=Functional Magnetic Resonance Imaging (Page 29); gCBF=Global Cerebral Blood Flow (Page 104); GM=Grey Matter (Page 6); HDL=High-Density Lipoprotein (Page 41); HR=Heart Rate (Page 69); IADL=Instrumental Activities of Daily Living (Page 18); ICA=Independent Component Analysis (Page 30); IMPACT=Improvement in Memory with Plasticity-based Adaptive Cognitive Training (Page 170); LDL=Low-Density Lipoprotein (Page 41); LEARN=Latent Early-life Associated Regulation (Page 10); MCI=Mild Cognitive Impairment (Page 17); MD=Mean Diffusivity (Page 35); MMSE=Mini Mental State Examination (Page 16); NFT=NeuroFibrillary Tangle (Page 7); NINCDS-ADRDA=National Institute of Neurological and Communicative Disorders and Stroke - Alzheimer’s Disease and Related Disorders Association (Page 23); NIRS=Near Infrared Spectroscopy (Page 105); NP=Neuritic Plaque (Page 7); PCC=Posterior Cingulate Cortex (Page 13); PA=Physical Activity (Page 62); PCR=Polymerase Chain Reaction (Page 131); PET=Positron Emission Tomography (Page 13); PHF=Paired Helical Filaments (Page 7); PiB=Pittsburgh Compound B (Page 13); pTAU=Phosphorylated Tau Protein (Page 7); rCBF=Regional Cerebral Blood Flow (Page 28); ROI=Region of Interest (Page 26); ROT=Reality Orientation Therapy (Page 95); RPE=Rate of Perceived Exertion (Page 112); RT=Response Time (Page 58); SEM=Standard Error of the Mean (Page 114-foo note); SPECT=Single Photon Emission Computed Tomography (Page 28); SPM=Statistical Parametric Mapping (Page 24); sSTAI=State-Trait Anxiety Inventory (State Anxiety Subscale) (Page 125); TAU=Tau Protein (Page 6); TBM=Tensor-Based Morphometry (Page 35); THT=Transient Hypofrontality Theory (Page 104); TIV=Total Intracranial Volume (Page 163); TMT=Trail-Making Test (Page 33); t-TAU=Total Tau Protein (Page 21); VBM=Voxel-Based Morphometry (Page 35); VLDL=Very-Low-Density Lipoprotein (Page 41); VO2max=Rate of Maximal Oxygen Consumption (Page 69); WCST=Wisconsin Card-Sorting Test (Page 69); WM=White Matter (Page 6).
1. NEUROMOLECULAR AND COGNITIVE BACKGROUND OF ALZHEIMER’S DISEASE AND MILD COGNITIVE IMPAIRMENT, WITH FOCUS ON THE ROLE PLAYED BY THE APOLIPOPROTEIN E GENE

GENERAL OVERVIEW AND OBJECTIVES

Alzheimer’s disease (AD) is the most common cause of dementia (Fratiglioni et al., 2000). It is a pathology that yields a drastic impact on a person from a neural, cognitive, psychiatric, emotional and functional perspective, with a huge impact on families and caregivers. Epidemiologic and statistical studies have suggested that the impact of AD on the future population will increase dramatically (Brookmeyer, Johnson, Ziegler-Graham, & Arrighi, 2007) and its economic burden will be a substantial issue (Rice et al., 2001). Such a devastating increase in AD costs, incidence and prevalence is paralleled by progress in AD-related research, but is far from being mirrored by scientific success: the cause for this disease remains at present unknown, and at present available treatments can only alleviate the cognitive symptoms. It is necessary to test new therapeutic pathways, in order to stabilise or improve the quality of life of present and future patients, but also to have a better understanding of the specific features of the disease that can be modified by behavioural activities. At the present time, engaging in physical exercise and in activities that stimulate cognitive functions are the most promising forms of non-pharmacological treatment. Many studies have indicated that these two therapeutic forms are beneficial, but the findings have been mixed, with some studies suggesting benefits and others instead reporting no effects. The main issue in this field of research is the lack of a strong rationale that may account for the specific mechanisms by which changes in behaviour would translate into meaningful structural and functional modifications to variables associated to AD.

The purpose of this dissertation is to contribute knowledge to this field of research and rethink non-pharmacological stimulation in the light of more fruitful paradigms and the availability of more sophisticated techniques to ascertain their efficacy. The studies so far published are a valuable database which needs to be reviewed with a critical eye, to identify the most appropriate descriptive framework of the disease that could be targeted by a treatment. For this reason Chapter 1 will include an extensive review of the main variables modified by AD. In
Chapter 2 the focus will be on physical exercise and engagement in cognitive activities: both literatures will be reviewed critically in order to classify the studies according to the different frameworks detailed in Chapter 1. Particular attention will be paid to the genotype for the Apolipoprotein E (ApoE) gene, which will be studied in relation to the features of AD pathology and in association to the effectiveness of both physical activity and cognitive stimulation. Aims and objectives of every single experiment will be described in Chapter 3, whereas the experimental part will be included in Chapters 4 and 5, followed by a general discussion.
1.1 Alzheimer’s Disease: Theories and Clinical Relevance

The clinical features of AD are based on the presence of particular molecules’ deposition in the central nervous system, accompanied by grey matter (GM), white matter (WM) and synaptic loss, brain functional deficits and behavioural decline. Various types of hypotheses have been suggested and tested (An et al., 2008), and most of them provide an explanation of the disease from a molecular perspective (Hardy & Selkoe, 2002; Pleckaityte, 2010). Numerous chemical alterations have been reported in neurotransmitter regulation (Lai, Ramirez, Tsang, & Francis, 2007), and this pattern appears to be linked to but independent from physical phenomena like loss of WM and synapses (Francis, 2005). Based on neurotransmitter’s deregulation, the main pharmacological treatment for AD at present is represented by medications that aim to restore the functionality of the cholinergic system (believed to be the cause of the underlying cognitive deficits). This paradigm has been recently flanked by the study of those polypeptides that are histological evidence of this disease: the Beta Amyloid peptide (Aβ) and the Tau protein (TAU). These are specific specimens that allow pathologists to diagnose AD with selectivity and specificity close to 100%. In fact, histological post-mortem investigation of the brain is usually the only method to diagnose AD. As autopsies cannot be routinely performed in the clinical setting for obvious reasons, recent research has put great effort in understanding the mechanisms of Aβ and TAU, trying to find alternative ways to visualise the molecules in vivo and diagnose AD prior to death. Unfortunately, histological assessment still remains today the only accurate methodology to diagnose presence or absence of AD, and the hotchpotch of available methods allows clinicians to come up only with a diagnosis of “probable AD”, albeit with a degree of validity and reliability that seems to be promisingly high.

1.1.1 Main Hypotheses

Aβ (Fagan & Holtzman, 2010) represents the final product of the chain of chemical reactions starting with the process of expression of the Amyloid Precursor Protein (APP) gene, located on chromosome 21. APP codes for the homonymous protein, which may undergo processes of cleavage by three different enzymes: α, β and γ secretase. Even if recent evidence has partially confuted this view, sequential cleavage by β and γ secretase at the two opposite terminals of the molecule is thought to occur only in the presence of pathological processes (Haass & Selkoe, 1993). This sequence of reactions transforms APP into Aβ. Aβ is a polypeptide that can be composed of 35 to 43 monomers, and the most common sequences consist of 40 (Aβ40) and 42 (Aβ42) elements. Aβ40 and Aβ42 possess different fibrillogenic properties with the longer
polymer tending to produce more fibrils and giving a bigger contribution to the deposition of what is considered the key hallmark of AD: the neuritic (or senile) plaques (NP). NP are extracellular depositions of Aβ that may locate anywhere in the brain and cerebrospinal fluid (CSF). Autopsy studies have revealed that NP may be mildly distributed (normally referred as “diffuse plaques”) and lacking any glial involvement that suggests neural insults, or focus on a particular region, forming deposits believed to have neurotoxic properties and triggering glial processes associated with neuronal damage (Selkoe, 1997). Since Aβ is a relevant molecule that drives the diagnosis of the pathology, it would be natural to infer that Aβ does cause AD. The “Aβ hypothesis” (Hardy & Selkoe, 2002) is supported by genetic evidence: all forms of familial AD (due to mutations of the APP gene, or Presenilin-1/Presenilin-2 genes, located respectively on chromosomes 14 and 1) are connected to alterations in the mechanisms that are involved in the production of Aβ from APP (Jack, Knopman, et al., 2010). Moreover, in Down syndrome, genetically described as trisomy of chromosome 21, the likelihood of developing AD is close to 100%, and this is interpreted as an amplified mechanism of APP gene expression (Lott & Head, 2001). Genetic evidence and constant presence of NP in post-mortem assays support the classic view that Aβ causes AD.

TAU protein (Medeiros, Baglietto-Vargas, & LaFerla, 2011) is coded by the homonymous gene (on chromosome 17), and belongs to the family of the Microtubules-associated Proteins; it is located inside the cells and its role is to stabilise axonal microtubules. Since axons have an important cytoskeletal component, long-distance communication between neurons seems to depend on the functionality of TAU. Along the chain of TAU amino acids there are many sites where a phosphate group can bind. A phosphorylation reaction normally activates or amplifies the function of a protein, and the phosphorylated version of TAU (p-TAU) rules the regulation of the neural cytoskeleton. When TAU hyperphosphorylates and self-aggregates into Paired-Helical Filaments (PHF), it appears to neurotoxically lead to a significant disruption of the microtubular regulation of the axons, compromising the correct synaptic communication and thus leading to cytoskeletal collapse that may cause cell death or synaptic loss. Deposits of p-TAU filaments normally take the shape of Neurofibrillary Tangles (NFT). The “TAU hypothesis” (Maccioni, Farias, Morales, & Navarrete, 2010) is mainly supported by the histological evidence that NFT involvement shows a gradual localisable progression that appears to be consistent in all AD cases, whereas this progression is not as visible for NP lesions (Braak & Braak, 1991). Since these findings have not been rejected, but also replicated (Braak & Braak, 1997), the alternative classic view posits that TAU hyperphosphorylation causes AD.

In brief, Aβ and TAU represent the most studied variables in this field of research: they permit the final diagnosis of the disease and can characterise the progression of the spectrum from
normality to dementia: Aβ mainly as an extracellular marker, with a possible amount depositing internally (WildBode et al., 1997), and TAU mainly as an intracellular marker, with a little amount of NFT located outside the neurons (the so-called "ghost tangles"). Both peptides display important connections with the cholinergic system: the nuclei which are responsible for the cholinergic innervation to the mediotemporal and limbic structures and to the associative areas of the neocortex are affected by NFT lesions (Mesulam, Shaw, Mash, & Weintraub, 2004). In addition, this neuronal population is influenced by Aβ, which is reported to be negatively associated with the number of cholinergic neurons in the basal forebrain (Auld, Kornecook, Bastianetto, & Quirion, 2002).

However, the classic view of either one of these two molecules (or both) defining a causal relation with AD pathology has been challenged by opposing evidence: the aforementioned lack of mutual correspondence between cognitive involvement and NP formation (Braak & Braak, 1991) does not support the Aβ hypothesis; and on the other hand, the loss of synaptic conjunction normally visible in AD appears to be independent from the presence of NFT (Cash et al., 2003). Further absence of support for the TAU hypothesis comes from the evidence that TAU is not specific to AD but is involved in many other pathologies of the brain. Whether NP and/or NFT are the cause or not of AD onset, they are always visible in AD pathology and thus represent two crucial biomarkers for this disease. Their systematic presence suggests that they might either be by-products of AD, or even part of mechanisms of defence against neural damage caused by the disease (Smith, Joseph, & Perry, 2000). Furthermore, Aβ (Smith, Casadesus, Joseph, & Perry, 2002) and TAU (Lee et al., 2005) have been reported to have antioxidant properties against oxidative stress, the neurotoxic action of free radicals (small molecules able to penetrate within the neuronal membrane), nitrogen and oxygen species (Jensen, 2003). From this opposite point of view, Aβ and TAU belong to the category of molecules involved in mechanisms of neural plasticity: in which the brain is subjected to the action of neurotoxic factors and the brain triggers the activity of protective factors as a response. Unfortunately there seems to be a limit to the efficiency of these molecules, and eventually the burden to make up for the effects of neurotoxic factors becomes too high. From this point of view, considering the protective nature of Aβ and TAU, AD onset can be defined as a progressive failure of neuroplasticity mechanisms (Mesulam, 1999, 2000). This idea of Mesulam’s is corroborated by evidence that NP and NFT deposition has been found postmortem (Price & Morris, 1999) and estimated in vivo (Small, Kepe, et al., 2006) in brains of non-demented individuals.
Fig. 1.1 The LEARn model hypothesises that early-age environmental factors accumulates contributing to the formation of an epitype co-moulded by genetics as well. A second hit on the epitype would force the system to cross the threshold of maximum burden (1.1a). At that point the mechanisms of neuroplasticity do not succeed in their task anymore and they start exhibiting neurotoxicity (1.1b). From Lahiri et al. (2008) and Mesulam (1999).

---


Lahiri et al. (2008) put another hypothesis forward, stressing the role played by environmental factors. Leaving aside all forms of AD that depend exclusively on genetic or environmental causes, most of the sporadic cases are reputed to be the result of a long-lasting process of environment-driven genes expression. The Latent Early-Life Associated Regulation (LEARn) model proposes that early-age environmental insults impact on genes, contributing to the formation of an “epitype”, a pattern of altered genetic expression (Lahiri & Maloney, 2010). This first early-age hit on the organism does not lead to any short-term neurotoxic effect. A second hit may occur later in life and the further altered genetic expression leads this time to abnormal gene regulation, which triggers the manifestations of the pathology. According to the authors, both hits are necessary, none of the two is sufficient: both must take place in order to provoke AD.

These two complimentary hypotheses (Figure 1.1) do not tackle a specific molecular mechanism but are more general and of wider ambition. They go beyond the identification of the correct aetiological factors (Lahiri & Maloney, 2010) but at the same time aim to create a satisfactory model that takes into account all relevant variables that might play a role in AD conversion. A linear cause-effect relation responsible for AD pathology is not needed (Terry et al., 1991), it is instead necessary to understand how to conciliate evidence coming from any step of the normality-to-AD continuum.

1.1.2 MAIN AD PEPTIDIC BIOMARKERS

1.1.2.1 POST-MORTEM EVIDENCE

As mentioned before, NP and NFT represent the two main biomarkers of AD and the only way to assess the impact of both peptides on the brain with precision is a histological assay of the central nervous system tissue. In a seminal study, Braak & Braak (1991) dissected 83 human brains mapping the progression of NFT and NP damage: they identified 6 stages of neurofibrillary change and 3 stages of plaques deposition: NFT involvement originates from the transentorhinal region (transentorhinal stages: 1 and 2), progressively involving the hippocampus together with mediotemporal and limbic areas (limbic stages: 3 and 4), and finally evolving into the neocortex (iso/neocortical stages: 5 and 6). Conversely, NP maps appeared to originate from basal portions of frontal, temporal and occipital areas in limited amount (stage 1), first multiplying and spreading everywhere in the cortex, with primary motor and primary sensory cortices excluded (stage 2), and subsequently involving also subcortical structures, with a significant increase in cortical density (stage 3). At more advanced stages,
the mediotemporal and limbic formations show a significantly lower density of NP involvement compared to the other lobes (Braak & Braak, 1991; Price & Morris, 1999). In these studies no effort was made to obtain a homogeneous sample and simple cognitive screening measures were taken to create a sample that could cover the whole spectrum of the disease. Arnold et al. (1991) focused on a sample of individuals who had been diagnosed with probable AD from 3 to 15 years prior to death and described a massive NFT involvement of mediotemporal and limbic structures, especially of the entorhinal cortex, subiculum and cornu ammonis 1 (CA1) hippocampal subfields. The remaining hippocampal areas were spared. Temporo-polar, perirhinal and posterior parahippocampal cortices were also reported to be heavily affected, whereas limbic and paralimbic structures displayed a high degree of variability. Cerebral subdivision into lobes displayed a preference of NP for the temporal cortex (the richest) and a relatively mild involvement of the limbic lobe (the poorest). The other lobes were classified in-between. From these three early studies it emerges how NFT deposition follows a precise pathway, whereas the distribution of NP displays more variability. A later study (Delacourte et al., 1999) confirmed this idea: autopsies were performed onto a set of brains with heterogeneous AD involvement and the latest stages of the pathology clearly displayed a massive NFT and NP involvement, but in the middle stages it was only neurofibrillary contribution that appeared to be always hierarchically well determined, whereas NP deposition seemed to be extremely variable. Part of this fluctuation is due to the enormous variability intrinsic to a diagnosis of probable AD when severity is not taken into consideration. A fourth autopsy study also provided corroborating evidence: Terry and colleagues (1991) came to the conclusion that the concentration of NP did not correlate with the level of dementia severity. NFT increase was reported to correlate modestly with cognition, but synaptic loss was the main correlate. This idea was corroborated by the results by Einstein and colleagues (1994) who studied the relationship between presence of plaques and neuritic structure of hippocampal granule cells in in vitro human post-mortem tissues (5 AD and 3 controls). No plaques were visible in controls, whereas they were well-defined in AD patients, allowing the authors to create a categorical variable (interplaque-plaque region); the neuritic structure was not different when comparing the two different areas, suggesting that plaques do not represent a key variable for axonal health. The authors concluded that the picture was that of deafferentation due to deprivation of inputs. As p-TAU normally controls and modifies the cytoskeleton associated with axons, it would appear logical to conclude that the hyperphosphorylation of this protein should account for synaptic loss. However, Cash et al. (2003) found that the reduction in synapses distinguishing AD from control brains is independent from PHF deposition, probably tending towards other mechanisms involving oxidative damage and mitochondrial dysfunction. In brief, evidence from autopsies confirms
that TAU and Aβ are two key-biomarkers of AD, but variability in distribution does not allow us to conclude that these peptides represent the actual cause of the disease.

Histological evidence shows a progressive affection of the brain originating from early-pathology key areas, evolving in time, and involving a big amount of cerebral tissue as the pathology reaches a remarkable level of severity. AD is normally suspected only when there is reasonable clinical impact on everyday life and cognitive function, but, as already mentioned, presence of AD biomarkers is visible also when cognitive difficulties are not yet detectable. Presence of initial cognitive deficits was reported in correspondence with the transentorhinal and limbic NFT stages (Mesulam et al., 2004). When the early phase of AD is completely asymptomatic, the comparison of an individual in the early stages of AD to a healthy adult does not display any gross behavioural degree of diversity. On the other hand, a difference is detected comparing the pattern of NFT and NP deposition: the progressive distribution of NFT and NP in the ageing process without dementia differs substantially: at least a minimum amount of NFT exists in 60 years old healthy brains, but NP may also be completely absent even at a very advanced age (Price & Morris, 1999). A community-based sample of individuals who did not meet any clinical criteria for cognitive impairment showed variable amount of plaques deposition (Aizenstein et al., 2008). The localisation of biomarkers is also different, with NFT preferring determined hippocampal structures and NP spreading instead diffusely throughout the cortex (Price & Morris, 1999). The theoretical frame provided by the idea of neuroplasticity loss (Mesulam, 1999) and the LEARn model (Lahiri et al., 2008) are capable of accounting for all this evidence. From the study by Price & Morris (1999) it emerges also that the ageing process could be sufficient to trigger an expansion of brain NFT involvement: below 75 years of age, brains do not show CA1 and periamygdalar involvement, while over 75 years old brains do; moreover many temporal cortical areas start becoming involved at old age exclusively. Completing the picture, young non-demented brains were more likely to have fewer plaques than their older counterpart. Further evidence showed that the ageing process does not look to be related to NFT/NP deposition in a causal way but rather as just a simple association (Braak & Braak, 1997).

**1.1.2.2 IN VIVO EVIDENCE**

The histological examination of brain tissue represents the gold standard to quantify the amount of neural damage caused by the biomarkers linked to AD pathology, but cannot be performed in vivo. Methodological efficiency is also limited because of the impossibility to carry out longitudinal studies. Recent innovative findings revealed that central deposition of Aβ plaques and p-TAU filaments can be visualised through the use of a radioactive compound,
created to be implemented in a Positron Emission Tomography (PET) scan procedure to trace its intracranial retention with a fair degree of invasiveness. Usually, the tracer is injected and the duration of the retention period indicates the map of cerebral deposition of AD-related biomarkers. The \([\text{N-methyl}^{6}\text{C}/^{11}\text{C}/^{3}\text{H}]\) 2-(4'-methylaminophenyl)-6-hydroxybenzothiazole radiolabeled Pittsburgh compound-B (PiB) is a useful tool to visualise the aggregation of A\(\beta\) in vivo with a PET scan, since it is derived from a dye used in histological and in vitro studies. The map drawn by this tracer and histological maps appear to overlap quite well: longer retention was reported in associative areas of AD patients with little or no involvement of WM and cerebellum (Klunk et al., 2004). Prefrontal, lateral temporal, lateral parietal, precuneus and posterior cingulate cortex (PCC) were reported to differentiate uptake duration in AD from controls (Jack, Jr. et al., 2008). These findings were confirmed by another team, who graded the involvement of different structures, and found that associative cortices and the striatum showed bigger retention than non-associative cortices such as M1, V1, S1, thalamus and brainstem, with little involvement of hippocampal structure and amygdala (Edison et al., 2007). Unfortunately, although showing the distribution of NP, PiB uptake does not seem to allow visualisation of NFT with similar degree of correspondence (Lockhart et al., 2007).

Ikonomovic et al. (2008) confirmed the previous findings and reported that the correlation between PiB-PET scanning and histological analysis of cerebral tissue performed post-mortem is considerably high \((r=0.58\) for A\(\beta_{40}\) and \(r=0.82\) for A\(\beta_{42}\)). Unfortunately there are sporadic cases in which particular polymorphic forms of A\(\beta\) do not bind selectively to PiB compounds, and a Braak-last-stage individual may even appear similar to a healthy control when scanned (Rosen et al., 2010) or can be generally described as showing AD-indistinct PiB retention (Jack, Jr. et al., 2008). According to Ikonomovic et al. (2012), the rare cases of false positives/negatives might be due to a borderline condition of an individual, the presence of cortical atrophy (binding to CSF amyloid may give extremely variable results), or the presence of diffuse, rather than neuritic plaques, frequently observed in the cerebellum, which are not detected by PiB (Ikonomovic et al., 2008).

\(^{18}\text{F}-(2-(1-\{6-\{2-\{^{18}\text{F}\}\}\} \text{fluoroethyl}) \text{(methyl) amino}\}-2-\text{naphthyl} \) ethylidene) \((^{18}\text{F}-\text{FDDNP})\) is a second compound used in PET developed to accurately reveal biomarker location inside the brain. The \(^{18}\text{F}-\text{FDDNP}\) binding potential can correctly separate AD from controls (being unspecific for NP and binding to general AD-related peptides, NFT included) but appears to be less sensitive than PiB compounds through the various severity stages of the pathology (Tolboom, Yaqub, et al., 2009). Recently, other compounds have become available: 4-[\(^{18}\text{F}\)-18] fluoro-N-[2-[4-(2-methoxyphenil)-1-piperazinyl] ethyl]-N-(2-pyridinil) benzamide \((^{18}\text{F}-\text{MPPF})\) has been recently introduced: acethylcholine apart, other neurotransmitter systems do display
alterations in AD, and serotonergic transmission from the dorsal raphe nucleus to hippocampal
CA1 subfield is the target of $^{18}$F-MPPF (Barrio, Kepe, Satyamurthy, Huang, & Small, 2008).
Synaptic “pruning” has been reported to be the main structural and functional feature of AD
(Einstein et al., 1994), and since this denervation affects the connections between basal nuclei
and hippocampus from one side (Auld et al., 2002), and connection between hippocampus and
associative cortices from the other hand (Villain et al., 2008), the idea of “hippocampal
disconnection” can be tested using $^{18}$F-MPPF radiotracer, as it specifically targets one of the
“connection sites”. Recently, further $^{18}$F-labeled tracers have been introduced, under the name
of Florbetaben, Florbetapir and Flutemetamol (Rowe & Villemagne, 2011). These tracers lead
to results that are extremely well-correlated to PiB-based compounds (Villemagne, Mulligan,
et al., 2012), but have increased half-life (Barthel et al., 2011), which facilitates the use of
these scans in a wider range of settings. Taking into consideration the shortcomings of these
tracers, most studies however report generally convincing evidence of applicability of
radioactive compounds as a diagnostic tool for prodromal AD. Visualisation in vivo of central
deposition of TAU is still in progress (Villemagne, Furumoto, et al., 2012) and will plausibly be
another milestone in this subfield of research.
1.2 MULTIDIMENSIONAL PROGRESSION OF AD PATHOLOGY

The picture so far described is the one of a pathology that attacks the central nervous system and causes neural and synaptic loss with progressive involvement. Since AD modifies the brain, changes are expected in cognition with a progressive picture that should roughly resemble the one so far described in terms of cognitive skills tapping those affected areas. From a behavioural point of view, three progressive phases of AD are normally recognised: the preclinical phase, the prodromal phase and the dementia phase (Jack, Knopman, et al., 2010). These three stages correspond approximately to the three stages originally described in the autopsy study published twenty years ago (Braak & Braak, 1991), but the degree of difference between a biological and a functional classification must not be underestimated: particularly, the middle phase of AD represents a critical moment in the clinical history of a patient as it is the time the symptoms become evident and a diagnosis may be eventually made with a huge impact on the person and their family. Neuropathological hallmarks and behavioural evidence are highly variable at that stage, and there is no discrete passage from absence to presence of pathology. There is a continuum from health to disease that makes the scenario very delicate.

Since AD is a diagnosis that is highly connected to the cognitive, psychiatric and psychosocial spheres, it is very important to understand where exactly an individual may be placed along the line that goes from normality to abnormality. Diagnostic tools are particularly necessary, together with any markers that might predict the evolution of the person in time at multiple levels.

Behavioural measures do not strictly depend only on the underlying neural substrate as there are many other well-known or still unknown variables that might play an intervening role in the relation between brain and cognition. This is why it is possible to find a neural correlate of AD even when a clear cognitive deficit is limited or even absent (the preclinical phase).

Borrowing the hypothesis of Mesulam (1999), it is conceivable that any neuroplastic processes may function in an effective way until a threshold level, after which they cannot make up for the neural damage any more. The timeline of neuroplasticity mechanisms is a second important continuum that compliments the progression of the pathology; there is a theoretical area on the timeline where these mechanisms do work effectively and nothing is detectable in the cognitive sphere. Any early detrimental effect of a neurodegenerative disorder (AD included) is absorbed and “fought” through processes linked to an effective neuroplasticity system. The subsequent portion of the timeline should display a more severe effect on the brain by the pathology (that is the nature of neurodegenerative disorders) and functional recovery continues until it reaches its limit. The excess of AD-related burden would result in a
failure of cerebral neuroplasticity, with gradual cognitive involvement. At the beginning this is very subtle, but then it becomes visible and detectable using classic standardised neuropsychological tools.

Technically, AD severity can be described only in terms of progression along the axes of the neuropathological features of the disease that cause neuronal and synaptic loss, and it does not allow us to draw any conclusion about the functional level of people. On the other hand, the correspondent cognitive phenotype of AD pathology is a condition of severe memory decline and eventually dementia. The literature shows that there is no perfect correspondence between biomarkers and cognitive levels, as people with a massive PiB retention could be functionally mint (Rosen et al., 2010) and no correlation has been found between PiB retention and the Mini Mental State Examination (MMSE) test (Jack, Jr. et al., 2008). The reason for this discrepancy can be traced to the middle phase of AD. Braak & Braak (1991) reported a certain degree of variability in Aβ deposition in the middle stage of the disease; as a consequence, individuals in the intermediate stage of the disease may display high or low degree of PiB retention (Jack, Jr. et al., 2008) and there are cases in which molecular biomarker levels represent no helpful tool in clinical understanding the functional stage of the underlying pathology. This “in-between” stage of AD exhibiting so much variability has received particular interest during the last 30 years, and a merely-cognitive approach has proven the most effective one to describe clinically individuals in this intermediate stage.

In clinical environments it is impossible to diagnose the presence of probable AD with confidence when the very first cognitive symptoms manifest, as they are very general and might be indicative of a broad spectre of underlying conditions (Winblad et al., 2004). To date, the symptom that has historically been reported as the first cognitive marker for AD-related decline is episodic memory decline (Blasi et al., 2009). Other studies have recently revealed that the decline of episodic memory is mirrored by a similar early decline in semantic memory mechanisms (Bocti, Whatmough, & Chertkow, 2004). A fine-grained analysis of neuropsychological data can reveal important diagnostic information: particular properties of semantic recall can be a better predictor of AD pathology compared to standard neuropsychological testing (Forbes-McKay, Ellis, Shanks, & Venneri, 2005). These specific early-stage cognitive difficulties reveal the presence of a third continuum from the absence to the presence of dementia, which correlates with the previous two (the continuum of AD and the continuum of plasticity).

It is really important in clinical practice to identify cognitively those people who are neither normal nor can they be classified as demented and thus can be located in the prodromal
phase. It is unfeasible and impractical to attempt to detect asymptomatic people in the preclinical stage of AD, and for this reason any effort should be concentrated on the very first signs of the prodromal phase. That is why during the last twenty years there has been an increasing interest in and an increased use of the concept of Mild Cognitive Impairment (MCI).
1.3 MCI AS A SNAPSHOT FROM NORMALITY TO DEMENTIA

1.3.1 MCI DEFINITION

The idea of MCI dates back to the late eighties and has been actively used as a clinical entity in the study of cognitive decline since. Petersen (2004) listed the basic criteria of inclusion: people are defined as amnestic MCI when they have subjective memory complaints, confirmed by impaired performance in neuropsychological tasks that tap memory processes, in which their performance is significantly worse than that expected from people of the same age and education. Diagnosis of dementia and impairments in daily life activities (ADL) must also be ruled out. Obviously, like any other differential diagnosis, any other clinical diagnosis that may account for the cognitive deficits has to be excluded. Non amnestic MCI is possible, when the deficit is restricted in one cognitive area other than memory. MCI may also be divided in single or multiple domain, according to the involvement of other cognitive difficulties apart from memory (or other function), but dementia must always be excluded (Winblad et al., 2004).

Ideally, pure amnestic single domain MCI individuals are quite rare (Kramer et al., 2006): there is indeed almost always some sort of comorbidity. Finally, although AD-dementia seems a common evolution of amnestic MCI, the subtype of a progressive MCI does not predict the cause of dementia the patient will suffer from: non-amnestic MCI may evolve to AD-dementia and amnestic MCI may evolve to dementia due other form of neurodegenerative disorder (Hussain, 2007).

Recently, there has been an attempt to revise these criteria, or at least to report the need for more flexibility at the moment of diagnosis, as other variables have been described that play an important role in the outcome of MCI. ADL have been suggested as one of the fluctuating aspects (Artero, Petersen, Touchon, & Ritchie, 2006), as there may be some subtle alterations in the instrumental activities of daily living (IADL), and particular attention on this aspect would permit a more accurate distinction between MCI and controls (Mariani et al., 2008). Some source of variability has also been found also when different operational definitions of the neuropsychological impairment have been used (Trittschuh et al., 2011); this is relevant because variability in testing material and criteria is normal in clinical practice. The “official” criteria have been also criticised because of statistical weakness, as there may be reduced sensitivity (Artero et al., 2006) and discrepancies in clarifying progression from MCI to dementia exist across various studies. However, generally speaking, these shortcomings are negligible when it comes to dealing with patients: a diagnosis of MCI represents a really helpful tool because in the diagnostic settings there is the necessity to identify potential AD patients.
Given the extremely practical role for which this diagnostic label has been put into practice, the original definition of MCI appears thus to be exclusively cognitive. During the last decade there has been an increasing interest in the neural substrate of MCI and in its level of AD-related pathological evidence. It has become useful to investigate the diagnostic efficiency of AD-linked structural and functional markers in relation to samples of people classified as MCI through pure cognitive criteria, to be able to obtain more accurate and predictive diagnostic groups, and in an attempt to reduce heterogeneity among research samples.

Even though MCI appears to picture a well-determined moment on the continuum from absence of cognitive deficits to proof of dementia, it does not estimate the condition of associated neural involvement, neither quantitatively (the amount of lesion) nor qualitatively (the underlying aetiology). The cause for a cognitive deficit in the absence of dementia may be diverse (DeCarli, 2003): vascular pathology or dementia with Lewy bodies may display a stage of transient MCI (Petersen, 2004), but also depression (Hansson et al., 2006), or other psychiatric, metabolic or traumatic (Winblad et al., 2004) factors can account for cognitive impairment. Indeed, a histological study found that brains with an ante-mortem Clinical Dementia Rating (CDR) score of 0.5 (thus generally classifiable as MCI) or 1-3 (usually identifying demented individuals) may display various underlying pathologies (Saito & Murayama, 2007). Possible presence of AD represents however the main cause of concern when MCI is diagnosed.

Since MCI and AD are theoretically independent but pragmatically strictly related, a key issue is the identification of AD pathology when the person is in the MCI phase; this passage from an objectively proven cognitive diagnosis to an attempt to identify the underlying aetiology requires the support of those variables that are relevant in the diagnosis of AD. MCI and AD may represent two different sides of the same coin where MCI is actually a cognitive transition state of the progression to AD-dementia, but they are completely independent when MCI does not progress to AD-dementia. This opens the doors to two important consequences evident in the literature: first, as there is no well-determined biological/neurocognitive marker which can predict conversion without mistakes, we are somehow forced to consider MCI as one big heterogeneous group in the experimental setting. Any association to a biological variable found in a general MCI group is the result of the heterogeneity of the sample, and large variability is therefore expected in those results. Second, since each MCI case either may or may not convert to AD, research has focused on the molecules that are believed to be the biomarkers and clinical evidence associated to AD pathology applied to a MCI group. Recently recommendations for additions to the classification criteria for MCI have included a measure
of neuroimaging, in order to estimate the presence of AD pathology and foresee a possible
degree of further conversion (Dubois et al., 2007).

When dealing with MCI patients, clinicians aim therefore to 1) identify the possible underlying
cause for the cognitive deficits and 2) predict the odds of conversion to a diagnosis of probable
AD-dementia. To succeed in these tasks there has been an increasing interest in investigating
the distribution of the main AD biomarkers not only in the brain, but also in a more peripheral
sector of the central nervous system, allowing for less invasive investigations and longitudinal
testing to be carried out. Longitudinal studies with MCI samples aim to detect or take
methodological advantage of the conversion rate from MCI to dementia (or from preclinical
AD to full-blown AD). The idea of MCI lacks a proper “progression” element. In theory it is
possible to register a decline of cognition with maintenance of a diagnosis of MCI, but to date,
longitudinal research on MCI samples has almost always had the intention to record either the
odds of conversion as a function of a target variable, or differences in a target variable
between converters and non-converters without any specific focus on “decliners-non-
converters MCI”. In fact, the decliners-non-converters look statistically (Riemenschneider et
al., 2002) and neuroimaging-wise (Pagani et al., 2010) similar to converters.

1.3.2 MCI AND AD-RELATED PROTEINS

1.3.2.1 CENTRAL DEPOSITION
Filtering the autopsy and PiB studies through the looking glass of a cognitive diagnosis, the MCI
phase could be roughly classified as the limbic stage of the pathology (Braak & Braak, 1991,
1997) with variable NP deposition (reflecting a variable PiB retention staining) and a well
determined pattern of mediotemporal and limbic NFT involvement. Caselli et al. (2010) found
exactly this pattern, as their MCI brains displayed no NP difference and enhanced NFT
involvement in the entorhinal cortex, the hippocampus and the temporal cortex, a perfect
picture of the doorway to a limbic involvement. More detailed data were published by
Guillozet et al. (2003), in an autopsy study investigating normal and MCI brains: MCI displayed
significantly more NFT in the fusiform gyrus and the mediotemporal lobe compared to
controls, but nothing could be concluded about Aβ because of a too-high data dispersion. The
interpretational key to this heterogeneity seems to be the presence of AD (and thus the
inexorable future conversion to dementia), as non-converters have cerebral and cerebrospinal
AD parameters similar to controls, and converters similar to AD individuals (Forsberg et al.,
Moreover, PiB data showed how MCI individuals with higher baseline retention are more likely to convert than MCI individuals with lower cerebral load of Aβ (Okello et al., 2009).

### 1.3.2.2 Peripheral Deposition

There is immediate contiguity from proximal extracellular space and CSF inside the brain; this is the rationale why analysis of peripheral concentrations of biomarkers could represent a good estimate of central parameters (Skoog et al., 1995). It is well-established that people who receive a diagnosis of AD display significant increases in levels of phosphorylated and unphosphorylated CSF TAU, and decreases in the levels of CSF Aβ (Fagan & Holtzman, 2010). Power analysis confirms how systematic these findings are (Lewczuk et al., 2004) but no agreement has been reached yet on normative age-adjusted biomarker data (Winblad et al., 2004). There is proof of correlation between CSF p-TAU and NFT cerebral aggregation (Burger et al., 2005) and between CSF Aβ and NP deposition measured post mortem (Strozyk, Blennow, White, & Launer, 2003) or visualised with PiB-PET although CSF biomarker specificity and CSF p-TAU sensitivity appear to be low at diagnosis (Clark et al., 2003). Ratio of Aβ40 and Aβ42 (Lewczuk et al., 2004) and combination of TAU and various Aβ isoforms (Welge et al., 2009) have also been reported to be valid markers able to separate AD from controls. Focusing on PET radioactive tracers, Aβ42 was reported to correlate with global 11C-PiB uptake in a sample that covered the whole AD spectrum, whereas total TAU (t-TAU) was intriguingly found to correlate with global 18F-FDDNP (Tolboom, van der Flier, Yaqub, Boellaard, et al., 2009). A similar trend of association between TAU concentration and 18F-FDDNP retention was reported also in a small sample of MCI (Small, Kepe, et al., 2006): an increase in metabolic parameters of the compound was reported in individuals with high autopsy-determined levels of a phosphorylated TAU isoform (not investigated in the previous Toolboom et al. study). Both PiB/FDDNP compounds and CSF biomarkers can help classify an individual with good accuracy in the category of the healthy controls or in that of probable-AD patients. The cross-use of these techniques, together with other diagnostic tools, may sensibly improve accuracy of the diagnosis. It is of particular interest to explore how these parameters are related to a diagnosis of MCI and how good they are to predict progression of cognitive symptoms to a condition of AD-dementia (Mitchell, 2009). In vivo visualisation of PET-compounds retention and peripheral levels of biomarkers could be a relevant method to detect the presence of a pathology hiding behind a “borderline” heterogeneous cognitive phenotype. Effort has been put in the attempt to find adequate CSF cut-off scores for use in clinical practice, however, as there is yet no agreement on guidelines, and since differences in cut off scores among studies may result in different scenarios of findings (Hansson et al., 2006), their use and sensitivity/sensibility will not be covered in this chapter.
Research has aimed to find variables that can indicate the presence of AD as early as possible in the timeline of the disease; predicting the conversion would mean being able to estimate the presence of AD and would permit the start of proper intervention immediately. Various epitopes of TAU phosphorylation have been described, most remarkably the ones with a phosphate group at position 181 (threonine), 199 (serine) and 231 (threonine). As these different isoforms of p-TAU were found to have similar significance levels in the comparison between AD and controls (Hampel, Buerger, et al., 2004), it is convenient to refer to them simply as p-TAU. When measured at baseline, neither CSF t-TAU (Forsberg et al., 2008), nor $A_{\beta 42}$ (de Leon et al., 2006; Forsberg et al., 2008) was found to differ between MCI and controls. On the other hand, in the research by Hampel and colleagues (2004) both polypeptides were found to significantly differentiate MCI from controls and MCI from AD, and de Leon et al. (2006) reported a p-TAU difference between the aforementioned groups. This lack of consistency could reflect the heterogeneity of the MCI diagnosis. In fact, when splitting the group according to progression of cognitive symptoms, differences were found between future converters and stable MCI, while stable MCI remained undistinguishable from controls (Herukka, Hallikainen, Soininen, & Pirttila, 2005; Riemenschneider et al., 2002). The same findings were reported by Hansson et al. (2006), who specifically investigated a stratified sample of MCI in order to rule out any possible bias in the recruitment of patients. Also p-TAU was found to differ cross-sectionally between MCI who later converted to AD-dementia and MCI who did not (Ewers et al., 2007; Herukka et al., 2005). Hampel et al. (2004) confirmed the efficacy of $A_{\beta 42}$ to predict conversion but failed to report a similar result for t-TAU. This lack of systematic evidence for the role of CSF t-TAU can be accounted for by the results by Sunderland and colleagues (1999): they followed a group of AD patients for over two years, reporting no longitudinal changes in CSF t-TAU. This finding suggests that there is no continuous gradual increase in TAU concentration, but rather an early sudden change which remains stable as the pathology progresses. This boost may take place way before conversion, making a significant longitudinal change impossible to detect. Similarly, CSF $A_{\beta 42}$ levels were reported to be visible even in the absence of cognitive deficit (Fagan et al., 2009), suggesting a parallel early increase of this peptide, which would be steeper in the preclinical stage of AD and therefore undetectable in the transition from the preclinical and the prodromal stage to later steps. Further longitudinal evidence confirms this idea: whilst a MCI-control comparison revealed significant test-retest differences in p-TAU levels after a one-year follow-up (de Leon et al., 2002), the same team found no test-retest differences between MCI patients and controls after two years of follow-up for CSF $A_{\beta 40}$, $A_{\beta 42}$ and p-TAU concentration (de Leon et al., 2006). Similarly, no longitudinal changes were reported for CSF t-TAU in a sample of AD
patients (Andreasen et al., 1998). The early nature of TAU and Aβ CSF changes is also supported by another study (Hansson et al., 2006).

1.3.3 MCI AND BRAIN STRUCTURE

It is worth considering that the central and peripheral values of TAU and Aβ might depend on the rate of GM and WM that is available for NFT and NP to aggregate: the National Institute of Neurological Disorders and Stroke-Alzheimer Disease and Related Disorders (NINCDS-ADRDA) criteria indicate measures of mediotemporal atrophy as supportive evidence for the diagnosis of “preclinical probable AD” in the presence of an MCI condition (Dubois et al., 2007). Atrophy rate is not sufficient to foresee the rate of progression of AD, but it is a good predictor (de Leon et al., 2006), and better than biomarker’s deposition (Smith, 2002), for the underlying cognitive deficits and, therefore, the diagnosis. In fact, whole brain atrophy displays significant differences when comparing converter to non-converter MCI, slowly-progressing to fast-progressing AD, and even healthy adults to healthy adults who have later converted to MCI (Jack et al., 2004). However, simple visual inspection of brain atrophy is a poor indicator of possible future conversion (Hansson et al., 2006). Objective measurement of annual rate of total brain and hippocampal volume loss is on the other hand a significant predictor of conversion from MCI to dementia (Erten-Lyons et al., 2006). Hippocampal and amygdalar volumes differ between a sample of MCI and a sample of AD (Luckhaus et al., 2010) and, focusing on the mediotemporal lobe, hippocampal atrophy was found to be a frequent feature of AD and MCI compared to controls (deLeon et al., 1997). In a longitudinal fashion, Venneri et al. (2011) reported the correlates of MCI-to-AD conversion to be GM losses in frontal, temporal, fusiform and lingual gyri on the left, and in temporal, parahippocampal and fusiform gyri on the right.

It is therefore relevant to understand the impact of CSF biomarkers also in relation to the degree of underlying cerebral atrophy. Degree of global ventricular expansion was found to be negatively correlated with CSF level of Aβ, whereas expansion of the posterior right horn was related to increase in p-TAU and the ratios between p/t-TAU and Aβ (Thompson et al., 2009). Correlational analysis in a sample covering the whole AD spectrum (controls included) revealed an association between whole brain atrophy and CSF t-TAU only, unreplicated within each diagnostic group (Sluimer et al., 2010). Further research got more into detail and investigated specific subgroups and brain regions: Thomann et al. (2009) reported that a mixed sample of healthy and AD participants at various levels of severity displayed significant negative
correlations between rates of CSF p/t-TAU and GM volumes in various frontal, temporal and parietal areas. Furthermore, specific focus on MCI revealed an association between p-TAU increase and increased rate of GM loss in the right hippocampus and right middle temporal gyrus without any significant results on t-TAU levels. Similarly, subgroups of MCI and mild AD patients displayed significant negative correlations between CSF p-TAU concentration and hippocampal atrophy coded through a scale from 0 to 5, with no significance in a normal subgroup (Henneman et al., 2009). These findings were not replicated by Fagan and colleagues (2009), who investigated CSF biomarkers and atrophy rate in individuals who had received a CDR score of 0, 0.5 and 1. Taking CDR=0.5 and 1 (mildly demented) individuals together, correlations were significant between total brain volume and CSF p/t-TAU (positively) and $\text{A}_\beta_{42}$ (negatively); moreover, CDR=0 (healthy) participants displayed a significant negative correlation between total brain volume and CSF $\text{A}_\beta_{42}$ only, without extension to $\text{A}_\beta_{40}$ or p/t-TAU, but the volume of the hippocampus was found to be correlated to none of the CSF parameters in none of the groups. Finally, Schoonenboom and colleagues (2008) reported a mixed set of results: no structural volume-CSF concentration correlations were found when samples of AD and controls were analysed separately; vice versa, correlations were significant when all participants were included in one big sample. Any discordance between these three studies may be caused by differences in sample choice or in methodologies, as in the first study “Statistical Parametric Mapping” (SPM) segmentation and voxel-by-voxel analysis were used, whereas in the second study “Functional Magnetic Resonance Imaging of the Brain Library Software” whole-brain volumes were based on a proportion estimated after “Freesurfer”-automatic assignment to a stochastic reference brain; in the third study, finally, hippocampal atrophy was rated visually in a scale from 0 to 4. In addition, in the study of Fagan et al. (2009) the gap between brain scan and lumbar puncture was very long (2 years). Aware of this lack of correspondence in the literature, the team of de Souza (2011) focused on hippocampal atrophy of a mixed AD-plus-MCI sample, and they found that even lateralisation may play an important role, because whereas hippocampal bilateral volume correlated with CSF p-TAU but not t-TAU, left hippocampal volume correlated with t-TAU but not p-TAU; atrophy did not correlate with CSF $\text{A}_\beta_{42}$. Taken together these findings suggest some sort of complex association between peripheral AD biomarkers and cerebral GM atrophy, but the pattern of such association is not clear. Considering how many details and variables may display variability in cross-sectional designs, it is convenient to focus on longitudinal studies in order to find more univocal conclusions.

Hampel and colleagues (2005) investigated the degree of change of hippocampal atrophy and CSF t- and p-TAU levels in a small AD sample, finding a positive correlation between increase in
atrophy and change in p-TAU only. No significant r scores were found for GM, WM and CSF total volume. A similar, albeit more modest, pattern of results was found in MCI (Fjell et al., 2010). However, Sluimer and colleagues (2010) did not find any significant correlation with whole brain volume across groups of AD, MCI and controls. From these studies it emerges that there is no strong well-determined trend of association between peripheral concentration of biomarkers and cerebral volume loss. The absence of any evidence in test-retest studies comparing CSF peptides and cerebral matter loss reflects the lack of test-retest change in CSF biomarker itself, as mentioned before. This conclusion may generally support the initial idea extrapolated from the review of the autopsy studies: the increase of the neuropathological features of AD reflects a mechanism of synaptic loss (Einstein et al., 1994; Terry et al., 1991), which does not directly entail any increase in total, GM or WM atrophy rate. However, the rate of hippocampal atrophy seems to be connected to the pathological picture of the early mediotemporal involvement described in autopsies (Braak & Braak, 1991). Furthermore, "classic" literature suggests that whole-brain atrophy is better at predicting cognitive difficulties rather than peripheral presence of Aβ and TAU.

Correlation does not imply causation, but if there were a causal relationship between brain volume loss and neuropathological markers of AD, it would be theoretically spurious: if CSF levels of biomarkers actually depended on levels of central deposition, it would make more sense to investigate the correlation between PET compounds retention and atrophy rate. Recent studies have tackled this correlation, but unfortunately are limited to Aβ, as binding rate to TAU is too low. Jack et al. (2010) transformed CSF Aβ42 values into an estimated cerebral load and divided their patients into two groups, according to the level of this burden; baseline group membership and baseline cerebral atrophy were reported to independently predict the future progression of MCI pathology (with conversion or no conversion) in the expected direction, without interacting. A longitudinal study (Driscoll et al., 2011) focused on healthy adults: structural scans were taken at baseline and after ten years, followed by a post-hoc 11C-PiB-PET scan as well. Rates of atrophy were reported to be dependent on the ageing process only, without any interaction or main effect played by the rate of Aβ deposition. The authors interpreted these findings giving two alternative explanations: whole-brain atrophy and Aβ deposition are independent processes (it would be in line with the evidence reviewed on peripheral markers and atrophy) or the mechanisms of "mutual-triggering" begin at a later stage of the disease. The literature seems to provide no evidence of any sort of regular trend of relation between presence of AD-related biomarkers in the central nervous system and rate of whole-brain volume. Using this statement as a starting point, Tosun and colleagues (2011) performed an "independent parallel component analysis" to investigate the regional rates of
PiB compound uptake and atrophy in parallel. In a sample of MCI, two components of correlation were found: first, Aβ deposition in mediotemporal regions, inferior parietal cortex, cuneus and precuneus on the left, together with bilateral PCC and right orbitofrontal cortex was correlated with atrophy in mediotemporal and limbic structures, inferior temporal cortex, frontal areas and cerebellum on the left, together with the PCC bilaterally; second, Aβ deposition in left inferior frontal, inferior parietal and superior parietal cortices with bilateral PCC, thalamus, cuneus and precuneus was correlated with atrophy in the left parahippocampal cortex and right mediotemporal structures, amygdala, middle orbitofrontal cortex and gyrus rectus. This detailed pattern of results sparks ultimate enlightenment on the complicated interrelationship between biomarkers and GM loss in AD-pathology and how cerebral lateralisation and the position on the disease timeline play a very important role, particularly in a heterogeneous condition such as MCI.

1.3.4 MCI AND BRAIN FUNCTION

1.3.4.1 PET STUDIES
A further source of information is represented by the level of functional activity of the brain, as synaptic loss is directly related to glucose metabolism. Functional neuroimaging shows that AD is related to hypoperfusion in posterior areas, namely the PCC, precuneus and mediotemporal areas (Kogure et al., 2000). Mosconi (2005) reviewed the findings of "standard" PET studies in AD: the 2-deoxy-2-(18F) fluoro-D-glucose (FDG) tracer revealed reduced AD-dependent uptake in frontal and temporal areas, together with the PCC and a vital role played by the hippocampal structures and amygdala. The findings on frontal uptake in AD have not been always replicated (Edison et al., 2007). Moreover, there seems to be a graded spatial progression in temporo-parietal hypometabolism, comparable to the progression of the pathology in autopsies: early studies came up with the finding that the earliest functional correlate of AD is represented by hypometabolism in the PCC (Minoshima et al., 1997). Nestor et al. (2003) challenged this view, stating that a voxel-by-voxel design might mask a prior hippocampal involvement: they compared AD, MCI and controls in mediotemporal and limbic Regions of Interest (ROI) and found that MCI and AD displayed no metabolic differences in the Papez Circuit, that is the architecture connecting the hippocampus to the PCC; however, there was a significant reduction in AD's amygdala and associative areas compared to MCI, suggesting a progressive hypometabolism that affects in succession mediotemporal and limbic areas. As support to this conclusion, El Fakhri et al. (2003) pointed out that the mediotemporal structures have small dimensions when compared to larger cortical areas. Based on this, a
relatively limited involvement of those structures might produce a major effect, possibly even more remarkable than a larger rate of neocortical atrophy. In a further study, Anchisi and colleagues (2005) studied various subgroups of MCI, describing a layout of functional differences: when compared to stable MCI, individuals destined to convert to AD exhibited decreased FDG uptake in the inferior parietal cortex and PCC bilaterally, in left dorsolateral frontal and parahippocampal cortices, and in the right hippocampus; moreover, only bilateral dorsolateral frontal differences were found between MCI non-converters and healthy controls. Similar subtle baseline differences were found in temporal and frontal areas of stable MCI compared to controls in another study (Drzezga et al., 2003). In this paper the comparison between non-converters and MCI patients who have later developed AD unveiled a very similar pattern to the one described by Anchisi and colleagues (2005), with reduced metabolism in the bilateral PCC and right precuneal, temporal and inferior parietal areas. Converters showed marked metabolic reduction in the following areas compared to controls: right inferior prefrontal cortex, left middle temporal cortex, the anterior portion of the insula and, bilaterally, PCC, inferior parietal cortex, hippocampal and parahippocampal regions. PCC and left temporal and frontal areas were reported to distinguish converters and non-converters in another study (Nobili et al., 2008). Finally, converters were found to have reduced FDG metabolism in bilateral PCC, left precuneus and left fusiform gyrus compared to controls (Pagani et al., 2010). Prospectively, Chetelat et al. (2003) reported greater right reduction in glucose uptake in newly-converted AD's temporo-parietal and PCC, compared to MCI who did not evolve; no difference was found between non-converters and controls. Vice versa, regional rates of cerebral metabolism were unable to differentiate between MCI and controls in another study, where just MCI-AD differences were found in all ROIs across the brain (Forsberg et al., 2008). Exceptions aside, the conversion to AD-dementia seems to be mirrored by an underlying reduced metabolism in posterior regions of the brain, especially parietal, temporal and limbic areas, with a frontal involvement that seems to occur at a later stage. The passage from mediotemporal to limbic involvement has been suggested as the functional onset for an AD diagnosis to be made (Nestor et al., 2003). Test-retest designs corroborate the idea that MCI-to-AD conversion is accompanied by further reductions in brain metabolic configuration: Drzezga et al. (2003) found underlying hypofunctionality of the middle frontal gyrus in both hemispheres, with reductions in right medial frontal cortex and left parietal and PCC. Conversely, stable MCI individuals (who plausibly do not have AD) exhibited only a slight reduction in another medial frontal cluster, when compared to controls. While Drzezga et al. (2003) compared converters and non-converters to a subgroup of controls, Fouquet et al. (2009) described differences in test-retest changes in metabolism directly between converters and non-converters. Both groups displayed a similar metabolic
reduction in agreement with earlier studies, but converters showed an additional decline in the medial cingulate and ventromedial prefrontal cortex. This study suggests that it may be the involvement of medial and prefrontal, rather than limbic, areas that could represent the passage to full-blown dementia.

The variability in glucose metabolism might be partially accounted for by the impact of other variables, such as the deposition of Aβ and TAU in the brain. The association between the uptake of the two tracers (FDG and PiB) was reported to be significant in AD’s temporal and parietal but not frontal associative cortex (Edison et al., 2007). Conversely, no evidence of correlation between glucose tracers was found in a mixed sample comprising controls, MCI and AD patients (Devanand et al., 2010). Comparing AD patients and healthy controls, an increase in CSF t-TAU was associated with bilateral metabolic decline in PCC and mediotemporal cortex, with parallel increase in occipital areas, frontal areas, cerebellum and basal ganglia (Haense et al., 2008). The correlation between FDG absorption and CSF p-TAU but not t-TAU was found significant in MCI (Fellgiebel et al., 2004).

1.3.4.2 SPECT STUDIES

Another technique that allows researchers to investigate functional activity of the brain consists of various tracer-related Single Photon Emission Computed Tomographies (SPECT): SPECT draws the pattern of regional blood perfusion of a working brain, but is not included within the group of investigation techniques that bring supportive evidence to the neuropsychological assessment to diagnose “probable AD” (Dubois et al., 2007); the reasons for this exclusion are generally low values of sensitivity and specificity for AD. As already mentioned, regional Cerebral Blood Flow (rCBF) was reported to be reduced in AD in mediotemporal areas, precuneus and PCC (Kogure et al., 2000) and reduced in AD in the PCC only, compared to MCI (Luckhaus et al., 2010). Huang and colleagues (2003) compared rCBF of MCI-to-AD converters to that of stable MCI and controls and reported a decrease in perfusion rate in bilateral parietal lobe paralleled by an increase in the bi-prefrontal cortex, with no differences between stable MCI and healthy controls. Hirao et al. (2005) replicated the findings on the parietal cortex only, with an additional reduction of blood supply in the precuneus bilaterally. A third study revealed a similar biparietal-bicuneal component, with the addition of the right temporal gyrus (Borroni et al., 2006). A larger sample was tested by Habert et al. (2011), who reported a baseline less intense perfusion in various portions of the right parietal cortex in converters compared to non-converters plus a decrease in the right hippocampus. The results collected in research with SPECT yields a pattern of changes in blood perfusion
similar to the picture drawn by FDG metabolism, with a progressive reduction in posterior areas and an interesting augmented supply in frontal regions.

Simultaneous use of SPECT and other techniques brought some further results in MCI: through the use of a mixed structural-functional design, El Fakhri et al. (2003) described a picture in which atrophy and reductions in blood perfusion seem to represent two independent aspects of the pathology: inputting both variables together in a single model, converters and stable MCI differed functionally in amygalar blood supply and structurally in anterior cingulate (ACC) volume, but both techniques converged independently towards an involvement of the basal forebrain and the PCC. Similarly, values of volume loss and perfusion in the hippocampus and amygdala (measured with perfusion-MRI this time) were found to be uncorrelated in AD and MCI (Luckhaus et al., 2010). As a final note, Okamura and colleagues (2002) found that a composite index made of rCBF and CSF t-TAU had a superior discriminating power than either single variable to classify an MCI individual as a future converter or not.

PET and SPECT techniques have also clarified the functional correlates of stable MCI: lacking a progressive decline, stable MCI suggests absence of AD; nevertheless, glucose uptake and blood supply appear capable of distinguishing stable MCI from controls: while some studies reported no significant p value (Chetelat et al., 2003; Huang et al., 2003), Anchisi et al. (2005) found reduced rates of frontal glucose consumption in stable MCI; similarly, Borroni et al. (2006) found differences in bilateral orbitofrontal cortex and superior temporal gyri. Nobili et al. (2008) described a decreased FDG uptake in stable MCI in structures including PCC, temporal and frontal cortex; finally, Hirao et al. (2005) found hypoperfusion in the PCC bilaterally and in the right caudate, with no evidence of frontal involvement.

1.3.4.3 FMRI STUDIES
Studies of functional Magnetic Resonance Imaging (fMRI) have been only recently systematically implemented in AD with experimental paradigms: resting-state fMRI variables have been investigated in relation to AD pathology and MCI diagnosis. The study of resting-state fMRI activation completes the described picture of cerebral activity at rest and introduces an important framework to interpret the set of results so far presented. It is possible to identify two major procedures to compute the pattern of resting-state Blood-Oxygen-Level-Dependent (BOLD) signal: regional coherence and functional connectivity. Measures of regional homogeneity compute the level of inner correlation of BOLD signal inside a pre-decided region (e.g. the hippocampus) using a statistical coefficient “of concordance”. This has been reported to be a good diagnostic marker for AD-related decline, being decreased in PCC and precuneus AD (He et al., 2007) and MCI (Bai et al., 2008) compared to controls. A
direct inter-group comparison was performed by Zhang et al. (2012), who found reduced homogeneity in medioprefrontal cortex, precuneus, PCC and left inferoparietal lobule for AD compared to MCI and controls; differences between MCI and controls were found in the left inferoparietal lobule only. No correction for atrophy was carried out in this study, though, unlike the former two studies. Functional connectivity represents on the other hand the computation of the co-activation pattern of BOLD signal between distant cerebral structures, and reflects the temporal synchronisation between the two areas during the complete echoplanar functional scan; an Independent Component Analysis (ICA) may be performed in an exploratory fashion identifying functional components that co-activate at rest. An alternative method is defining one or more pre-determined ROIs (based on a hypothesis) and computing the map of correlation every other ROI or voxel displays with the seed. These two methods were compared recently and the use of ROIs and their correlations was reported to be more powerful to correctly discriminate an AD individual from controls, but the two techniques must be interpreted carefully in order to correctly classify an MCI individual, as none of the two was sufficiently powerful and only displayed a trend of disrupted connectivity (Koch et al., 2012). Moreover, selecting another seed region can lead to two completely different sets of results and for this reason it must be chosen with caution (Bedenbender et al., 2011). Elevated co-activations between distant structures in healthy adults are supposed to “hide” a functional circuit; a pathological reduction in this value would suggest a partial disconnection of such network. Several resting-state networks have been estimated, most of them are related to primary sensory/motor computation, and primary attentional/executive processes (Sorg et al., 2007; Trachtenberg, Filippini, Ebmeier, et al., 2012). Liu et al. (2008) reviewed the literature on whole-brain and ROI functional connectivity in AD: evidence shows a pathology-due disruption respectively in the overall correlational map and specifically between the hippocampus and a big set of anterior and posterior areas. A similar report stated that a circuit comprising left hippocampus, parahippocampal region, precuneus, gyrus rectus, dorsal cingulate and ACC is significantly impaired in AD compared to controls (Sheline, Raichle, et al., 2010). The idea of functional disconnection between distant areas inside the AD-brain has been supported by structural evidence affecting the WM that links mediotemporal areas to the PCC. The co-occurrence of hippocampal atrophy and hypometabolism of the PCC can be explained as significant disruption of the WM tracts connecting these two distant regions (Villain et al., 2008). The picture of progressive hypofunction of brain areas together with a pattern of hippocampal and, more generally, neuronal disconnection (in addition to the cholinergic denervation from the basal forebrain) can be accounted for with a particularly good degree of fitness by the idea of disruption of a set of circuits called Default Mode Network (DMN), composed by functional hubs located in the anterior and the posterior half of the brain.
PET, SPECT and fMRI studies have shown how some brain areas normally activate both when the participant takes part in a task-free brain scan and during the breaks between task trials (task deactivation). These hubs include the ACC and the medial prefrontal cortex in the frontal lobe, temporal and mediotemporal areas, the PCC and the precuneus, and the inferior parietal cortex (Buckner, Andrews-Hanna, & Schacter, 2008). When the brain does not engage in a specific task, it is not in an idle state, though: significant cognitive processes might however take place in an uncontrolled (but statistically irrelevant) way. Attendance to external stimulation, autobiographical memory, envisioning of the future, inner decision making and use of the theory of mind are examples of mental processes that might underlie brain activity at rest (Andrews-Hanna, Reidler, Sepulcre, Poulin, & Buckner, 2010). The idea of DMN has been extremely helpful in the organisation and interpretation of neurocognitive results in AD research, and at the same time it is a relatively simple looking glass to interpret brain function in AD. However, Morcom & Fletcher (2007) published a paper that supports a view of non-relevance of the DMN: they criticise the significance of this construct on theoretical grounds, bringing physiological evidence in support of their view. The idea that emerges from their paper is that the use of the DMN notion assumes really basic and still not well-clarified neuroscience concepts. Such diatribes about these basic concepts will inevitably translate into critics against the use of the DMN. However, despite this, the DMN has been massively studied in relation to AD and MCI.

AD pathology has been formally described in terms of disruption of part of the DMN (Greicius, Srivastava, Reiss, & Menon, 2004), and all previously mentioned PET and SPECT evidence fits this suggested model with precision, showing how its posterior component, albeit being more robust than the anterior component in terms of individual fluctuations (Kim & Lee, 2011), appears to be systematically impaired as the patients progress from the preclinical phase to dementia. Recent publications that have studied AD and MCI specifically in relation to the DMN using fMRI bring further evidence in support of this view. Using ICA, Qi and coworkers (2010) calculated the best-fit component for the DMN and reported a connectivity reduction in amnestic MCI in the left fusiform gyrus, right inferior parietal lobule and in precuneus/PCC on both sides; this diminishment was paralleled by an increase of connectivity in frontal clusters. Another exploratory study compared MCI to controls and reported decreased connectivity in parietal patches of the DMN with additional involvement of the left PCC and the right medioprefrontal cortex, but without any volumetric changes of these areas (Sorg et al., 2007). Rombouts et al. (2005) studied task deactivation as a measure of DMN in controls, MCI and AD, and found variations in brain function which depended on the task (episodic or working memory). The pattern resembled the DMN with both tasks but based on these task-dependent
discrepancies, task deactivation measures do not appear to represent the gold standard to study the DMN. ICA was recently performed by Jin et al. (2012) on amnestic MCI and controls in order to compare the DMN between groups in terms of functional connectivity. Decreased activity was found in left frontal gyrus, left prefrontal cortex, left middle temporal gyrus and mediotemporal structures, right angular gyrus and bilateral PCC and precuneus, with a parallel increase in left inferior parietal cortex, medial prefrontal cortex and middle cingulate. Irrespective of the specific goal to interpret the task-free functional scans as a function of the DMN, the only recurrent result is the involvement of the precuneus and the PCC as core areas of difference between controls and AD patients at various stages of the disease, an idea that perfectly fits the finding of WM dismantling between mediotemporal lobe and limbic regions. Sheline and colleagues (2010) compared functional connectivity of normal individuals with relevant PiB response to controls without a significant PiB staining, and found a decrease in the correlational map in a fashion similar to an “ordinary” AD-controls comparison, suggesting an association between Aß central deposition and brain function.

1.3.5 MCI AND COGNITIVE FUNCTION

In the diagnostic setting, the principal point is the identification of the underlying reason for the cognitive impairment. The use of structural and functional neuroimaging, together with the measurement of biomarkers obtained through spinal tap helps clarify the overall pattern of clinical evidence that can be translated in an accurate “probable” diagnosis. At the same time, marker information, together with their complicated pattern of interactions and with the measurement of their progressive change, may provide the individual with expectation of further future cognitive and functional decline and, eventually, possible progression to dementia. Nonetheless, the ultimate way to diagnose MCI is by neuropsychological testing: scores on cognitive tests describe the level of cognitive functioning a person is exhibiting and the difficulties they can have in everyday life. For this reason, it is necessary to detail how cognitive scores interact with AD pathological hallmarks, especially in the prodromal phase.

It is difficult to describe a general pattern of association between AD-relevant polypeptides and cognition: in a single sample of AD patients, Edison et al. (2007) found that memory recognition measures significantly correlated with whole-brain and associative-regional PiB distribution. Similar findings were obtained by Forsberg et al. (2008). In another study (Tolboom, van der Flier, Yaqub, Koene, et al., 2009), temporal PiB retention was found to be negatively associated with immediate memory recall only in MCI, whereas MMSE and memory
performance (both immediate and delayed recall) were associated with frontal PiB hyper-retention, and parietal metabolism was correlated with performance in the Trail-Making Test (TMT), the Stroop Colour Word Interference test, the Rey Figure (copy and delay) and semantic fluency. This team also reported associations between frontal $^{18}$F-FDDNP uptake and measures of immediate memory only. Regarding the association between peripheral biomarkers and MMSE, there is evidence for an association between baseline p-TAU (and not t-TAU) and MMSE drop in MCI (Buerger et al., 2002), but no correlation between baseline CSF p/t-TAU levels and MMSE decline in full-blown AD (Haense et al., 2008).

Attempting to predict the conversion from MCI to AD involves finding an association between a predictor and the cognitive interface of the brain. The evolution to dementia corresponds by definition to a decrease in multiple cognitive aspects and this drop appears to correlate with reduction of regional cerebral metabolism. Arnaiz et al. (2001) focused on three temporoparietal ROIs: one at the level of the basal ganglia and two respectively 13mm above and below the basal ganglia. A stepwise logistic regression revealed that the model with the highest predicting power for MCI-to-AD conversion was obtained with the left upper ROI and the combined score in the Digit Symbol Modalities Test (DSMT) and the TMT–part B, two measures of separate attention, attentional shifts, executive control and speed of processing. According to this team, the predicting power of episodic memory is nil, because it had already declined at the entrance to the MCI stage (amnestic MCI is indeed the most frequent subtype) and did not have any discriminating power any more within the MCI stage. Positive correlations were detected between temporal FDG metabolism and MMSE, category fluency, immediate recall and delayed recognition (simpler than recall and therefore less likely to display a floor effect) in AD (Edison et al., 2007). Fine-grained subdivision of MCI into two subgroups having respectively lower and higher learning scores (measured with the California Verbal Learning Test) revealed a reduced FDG metabolism in inferior parietal cortex, PCC and precuneus bilaterally in those who scored less than 7 at the test; vice versa, a metabolic reduction was visible in the right dorsolateral frontal cortex of those who scored 7 or more, compared to controls (Anchisi et al., 2005). This frontal component has been detected by other authors as well: according to the team of Schönknecht (2009), episodic recall performance in MCI individuals is correlated with notable glucose metabolism reallocation at rest in the frontal lobe. This over-recruitment of resources in the anterior sections of the brain is accompanied by an impaired cognitive performance and might be explained in terms of compensatory mechanisms (and thus possible use of extra executive skills). This finding resembles the pattern of conversion-dependent perfusion rate reported by Huang et al. (2003) with SPECT. These authors also investigated the pattern of correlations between brain functionality and
neurocognitive performance, finding negative correlations between a measure of learning (Word Paired Associates test) and left medial frontal blood flow, and between the DSMT and perfusion in the right cerebellum. The same conclusion was drawn by Luckhaus et al. (2010) who described hippocampal and limbic volumes being negatively correlated with frontal perfusion in MCI. These authors all explained this anterior component as a mechanism that compensates for posterior damage. Increased activity in the anterior hubs of the DMN supports this hypothesis as well (Jin et al., 2012). Based on this evidence, reallocation of resources appears to be a complex pattern that involves structural, functional and cognitive variables. Moving one step forward, a canonical correlation analysis was carried out on neural components (SPECT measures calculated with a previous principal component analysis) and neuropsychological tests together, revealing that a combined baseline pre-conversion predictor may have more predictive power than a single neural or neuropsychological variable (Borroni et al., 2006).

A notion that is for some aspects the antithesis to the idea of reallocation of resources is the concept of diaschisis, which may account for some of the co-variability found between two distant sources of evidence, such as the hypometabolism found in associative areas and the hippocampal volume loss. Villain et al. (2008) found a significant correlation in early AD between GM loss in the hippocampus and hypometabolism in the PCC/retrosplenial cortex, together with additional WM loss in the temporal lobe, cingular bundle, fornix and perforant path. This supports the idea that parallel structural and functional dysfunctions in distant areas may occur because of underlying WM damage. Smith (2002) reviewed the "hippocampal disconnection" hypothesis of AD formulated in the mid-eighties, which precisely indicates how the clinical changes of AD originate from the isolation of the hippocampus from associative areas of the neocortex. This idea is also supported by pure FDG-PET studies (Anchisi et al., 2005; Forsberg et al., 2008), by histological findings (Einstein et al., 1994; Terry et al., 1991), and appears to be clearly dependent on the architecture of WM structures.

As a final remark to complete the picture, there is also evidence of association between cognitive performance and cerebral volume: a recent study (Venneri, Gorgoglione, et al., 2011) found that a sample of MCI scored significantly worse than healthy participants in the paired associates and category fluency tests (testing respectively episodic recall and semantic access); based on this, voxel-based correlations extended to the two conjoined subgroups revealed that worse performance in the paired associates test was associated with GM reductions in the parahippocampal gyrus and superior temporal cortex bilaterally, in middle temporal and lingual gyri and the cerebellum on the left, and in the inferior temporal gyrus on the right; worse performance in the category fluency task was correlated to GM reduction in middle and
superior temporal gyri bilaterally, right uncus and left inferior temporal gyrus. In the same study, a bidimensional (anatomic + neurocognitive) correlate for MCI-to-AD conversion was detected as a combined low performance in episodic and semantic memory together with atrophy in a left temporal cluster.

1.3.6 MCI AND OTHER INVESTIGATIONAL TECHNIQUES

Many other neuroimaging techniques like for example Diffusion Tensor Imaging (DTI), Electroencephalography (EEG) or Tensor-Based Morphometry (TBM) have been used in the quest for further AD and MCI neural markers. In brief, the pattern of findings emerging from the use of any of these three techniques is similar to the picture drawn so far. DTI is a high-resolution MRI technique that measures indices of water diffusion in tissues; these measures have been demonstrated to be sensitive to integrity of WM structure, and can reveal subtle structural alterations in absence of volumetric changes. Parameters of Mean Apparent Diffusion (MD) and Fractional Anisotropy (FA) indicate respectively the average diffusivity and the directionality of diffusion. Furthermore, FA can also be split into a radial and an axial component according to the direction of the vector. Generally, research has revealed that AD and MCI patients differ from controls as they display decreased FA and increased MD in a large set of WM regions (Sexton, Kalu, Filippini, Mackay, & Ebmeier, 2011). TBM is a methodology of analysis of structural MRI data, similar to Voxel-Based Morphometry (VBM) as it estimates voxel-by-voxel differences between a brain and a “normal” template. The difference lies in the nature of the comparison: whereas VBM gives each voxel a probability score of membership in each of the three segmentation classes (WM, GM, or CSF) and compares the three submaps to a standard template, TBM is a deformation-based approach that looks at brain shape rather than brain composition: a standard-deformation method aims to estimate a global parameter of how different a brain’s shape is from a standard template; TBM focuses on a local perspective, and is based on a voxel-by-voxel extraction of a value called Jacobian Determinant, that indicates the difference in volume between that voxel and its neighbouring areas of a “normal” template (Ashburner & Friston, 2000). TBM revealed that AD and MCI individuals display atrophy of the hippocampus and the temporal lobe, plus an enlargement of the lateral ventricles, albeit for MCI a milder pattern was evident (Hua et al., 2008). Compared to fMRI and the other techniques, EEG has a better temporal resolution and provides with complimentary information about brain function in AD pathology. Resting-state EEG recordings display a normal ageing effect in the amplitude of the posterior cortical areas:
components of higher amplitude are taken over by slower waves; in AD this pattern is exacerbated and MCI shows once again an overall picture of transition (Babiloni et al., 2010). Other CSF molecules have also been examined, but for most of these there is no large number of papers that have validated their use (Tarawneh & Holtzman, 2010) and therefore will not be reviewed in detail.

1.3.7 MCI AND THE TIMELINE OF BIOMARKERS

All things considered, there is evidence for an ample set of AD biomarkers: CSF, functional, structural and cognitive variables which possess their own timeline of AD-progression and concur to draw an overall continuum that describes a general timeline of the disease. To the initial hypotheses suggested by Mesulam and Lahiri, it is possible to link a third model that can be drawn adjacent, which tries to put together all these biomarkers along a temporal dimension. According to Jack et al. (2010) there is necessity to pinpoint the temporal integration of the relevant biomarkers, correlating clinical evidence, histological biomarkers and neuroimaging data in a dynamic model that describes a neurophysiological cascade. According to the model by Jack et al., Aß plaques formation and NFT accumulation represent the beginning of this cascade, and are followed by a “lag-stage” in which the biomarking pattern change is silent. At this point, AD course is asymptomatic and only the polypeptidic hallmarks are visible. The second step is represented by alterations in processes connected to brain structure and functionality. The cognitive aspects represent the bottom of the cascade and the set of terminal symptoms that normally warrants a clinical inspection. This model allows us to interpret apparent ambiguities and to put more paradigms of research together in order to conciliate evidence coming from different sources of information: first, the idea of lag-stage would be coherent with the hypothesis of progressive failure of plasticity mechanisms (Mesulam, 1999). The evidence of increase in biological markers in absence of any relevant clinical sign of abnormality (Price & Morris, 1999) is coherent with a view in which the system generally keeps up with the overall neural change; early Aß and TAU involvement would be the mechanism by which the cerebral system successfully makes up for the damage caused by progressive failure of neuroplasticity. The end of the lag stage is the moment when the “point of no return” is crossed and the mechanisms are no longer beneficial but they irreversibly start contributing to the neurotoxic process that assails the brain. Second, Cash et al. (2003) pointed out that loss of efficiency in synaptic communication occurs before any significant change in TAU levels; according to this, synaptic damage could represent the trigger that provokes the early asymptomatic TAU and Aß changes. As a support to this view, Sunderland and colleagues
(1999) reported that multifaceted cognitive decline in AD is not accompanied by a similar t-
TAU increase in CSF. As further flanker support, levels of PiB retention were not described to
evolve with the progression of the disease (Rowe et al., 2007). This suggests that a steep
change in the two main biomarkers does not occur in the late stage of the disease, but just at
the beginning. The challenge between A\beta-first and TAU-first supporters is linked to the classic
view of either molecule as playing a causal role for the onset of AD. There is conflicting
evidence that alterations in both A\beta (Hardy, Duff, Hardy, Perez-Tur, & Hutton, 1998) and TAU
(de Leon et al., 2006) mechanisms precede any alterations of the other respective peptide.
Further evidence strikes a balance, suggesting that probably a NFT involvement occurs before
the first relevant depositions of A\beta, but is not necessarily associated to AD (Price & Morris,
1999). Placing the A\beta and TAU curves on the timeline of AD becomes vital when using
measures of ratio between the two peptides.

The model by Jack et al. predicts that brain atrophy and hypo-functionality occur at a second
stage. A gradual functional decline was described by Pagani et al. (2010), who found peculiar
areas of difference between controls, stable MCI, declining MCI (referred as "late converters")
and converting MCI. Particularly, according to Nobili et al. (2008)’s speculation, onset of frontal
hypofunction has been linked to IADL impairment and thus to a diagnosis of dementia due to
failing compensatory mechanisms associated with frontal areas. Luckhaus and colleagues
(2010) suggested that the rate of progression of functional and structural aspects in limbic and
mediotemporal regions does not follow the same ontogenetic line; hypoperfusion would start
in the preclinical stage and reach an asymptote in prodromal AD, whereas atrophy would
continue linearly. Jin et al. (2012) found functional differences between MCI and controls in
the absence of any volumetric differences in the medial temporal lobe, supporting the idea
that functional changes may be visible without any structural changes. Finally, atrophy was
reported in individuals with CDR scores of 0, having thus no objective cognitive impairment
(Fagan et al., 2009) and suggesting that neuropsychological functions are affected at the latest
stage of the model. As a general statement, structural decline starts before the onset of
neuropsychological symptoms (Smith et al., 2007).

Within the latest stage of this timeline, cognition would be affected. The idea of a gradual
involvement of different cognitive functions is intrinsic for the concept of MCI (Petersen,
2004). However, prior to any impairment objectively confirmed with routine
neuropsychological tests, subtle difficulties may be detected (Forbes-McKay et al., 2005). The
concept of Subjective MCI has been suggested in the presence of objectively unconfirmed
cognitive complaints (Reisberg & Gauthier, 2008) and it appears to have a specific
intermediate neural correlate between controls and MCI individuals (Saykin et al., 2006).
Neuropathological and behavioural evidence shows that there is no discrete passage from absence to presence of AD. As it is really important to understand where exactly an individual may be placed along the line that goes from normality to abnormality, diagnostic tools are particularly necessary, together with any marker that can predict the evolution of the person in time at multiple levels.

Although it does have methodological shortcomings, the idea of MCI is really helpful from a clinical perspective, as it identifies potential AD-affected individuals, who may start undergoing a process of treatment and prevention that would be impossible if such a practical diagnostic entity did not exist.
1.4 THE ROLE OF THE APOLIPOPROTEIN E GENOTYPE FROM HEALTH TO PATHOLOGY

Unfortunately, while no effective treatment for AD has yet been discovered, there has been an effort to identify important variables that play relevant mediation roles to the onset of the disease. Finding these variables could permit further enlightenment on the molecular mechanisms underpinning the pathology and it would have repercussions in preventing the disease. These entities are often referred to as risk or protective factors, since they have been reported to be associated with an increase or decrease in the likelihood of developing AD. These variables are not meant to be described as part of a causal relation with AD, but rather it is the intent to seek correlational associations with the disease or significant predictors that modulate the odds of AD onset and/or its progression. For convenience, the absence of a risk factor is also considered as a protective situation and vice versa. These variables may be either genetic or non-genetic. Non-genetic variables are often named as environmental. Environmental variables do not necessarily deal with the external environment, but they identify biological variables that are simply non-genetic. Diet choice and stress response are two examples of environmental factors: both refer to external factors of everyday life but their influence on AD dynamics relate to the effect these variables have on biological compounds inside the human body (in this case metabolism associated to nutrition and hormonal stress response) which are known or believed to directly or indirectly affect the nervous system. Genetic factors are of particular interest from this point of view: aside from giving risk or protection for a disease, they are categorical variables with a limited number of levels, which lend themselves to simple statistical procedures and results interpretation. If there is interest in a single allele, genotypes can be converted into a binomial variable, according to the presence or the absence of the allele, excluding undesired genotypes or assuming that the presence of any other allele is controlled for or irrelevant.

1.4.1 THE APOE GENE

The ApoE gene has been extensively studied in relation to AD during the last 20 years. Epidemiological research shows that there are three main isoforms for this gene: $\varepsilon_2$, $\varepsilon_3$ and $\varepsilon_4$. A review of allelic frequency distribution in the world showed that $\varepsilon_3$ allele is the most common one, with a frequency between 53% and 89% with value shifts due to ethnic (Corbo & Scacchi, 1999) or geographic (Eisenberg, Kuzawa, & Hayes, 2010) factors. $\varepsilon_2$ and $\varepsilon_4$ alleles are
relatively less common (usually around 10%) and further isoforms such as \( \varepsilon_1 \), \( \varepsilon_5 \) and \( \varepsilon_7 \) have been described as very rare (Maeda et al., 1989; Ordovalas, Litwackklein, Wilson, Schaefer, & Schaefer, 1987) and, together with \( \varepsilon_2 \), will be not considered here\(^4\). Based on these data, \( \varepsilon_4 \) and \( \varepsilon_2 \) carriers (homozygotes and heterozygotes) are estimated to be respectively around 25% and 15% of most of the world’s population. \( \varepsilon_3 \) has been usually reported as the standard allele, whereas \( \varepsilon_2 \) and \( \varepsilon_4 \) are alternative isoforms connected to altered prevalence of a huge number of cardiovascular, viral and neurodegenerative pathologies (Hill, Lhattacharjee, & Neumann, 2007).

Focusing on AD, \( \varepsilon_4 \) allele is a well-established risk factor for sporadic late-onset forms of the disease (Strittmatter & Roses, 1995), and is associated with younger age of onset for AD (Sando et al., 2008) and progressive MCI (Herukka et al., 2007). In most of the papers cited so far the ApoE genotype was investigated and used as a predictor or categorical covariate, but none of the articles reported it as a variable with a significant impact on the outcome values. This appears to be a contradictory conclusion, as an effect of genotype should be expected on variables involved in AD progression. However, it is necessary to look at the literature that has specifically investigated the effect of the ApoE makeup, because subtle methodological or statistical reasons can mask the effect of a variable when proper procedures of analysis are not implemented. In many instances the \( \varepsilon_4 \) allele has been treated as a binomial variable (presence or absence), without specific focus on homozygosis. Moreover, research has tried to isolate the effect of the ApoE genotype on AD-related variables from that of ageing process and AD-associated biomarkers: as such a substantial body of studies recruited a sample of young adults, on whom the neuropathological cascade of the disease previously described would not

\(^4\) The ApoE genotype can be determined either by analysis of the DNA sequence or by analysis of the direct phenotype and thus the ApoE protein. Most of the times a classic Polymerase Chain Reaction (PCR) has been carried out, with the amplification of a short fragment containing the loci of the two main nucleotide changes and the digestion of this fragment by enzyme restriction cutting the sequence in correspondence of the configuration of the base pairs. Old papers reported the use of a single enzyme (Hha I or its isoschizomers), later criticised for producing too many fragments that made the interpretation more difficult under UV light. A more novel approach has been recently utilised with two enzymes (Afl III and Hae II) which create a configuration of fragment lengths that is easier to interpret. A real-time PCR machine has been sometimes used and sometimes the proteins were assayed. The differences in genotyping methods are not expected to alter the results, as each of those is directed to the correct identification of the three main isoforms \( \varepsilon_2 \), \( \varepsilon_3 \), and \( \varepsilon_4 \). An important issue could be the detection of the remaining alleles (\( \varepsilon_1 \), \( \varepsilon_5 \) and \( \varepsilon_7 \)), as the standard PCR methodology does not allow a biologist to detect any polymorphisms that are located outside the amplified string. As an example, Yamanouchi et al. (2001) described the \( \varepsilon_5 \) isoform as a 24-base-pair (thus 8 amino acids) insertion detectable only by amplifying and sequencing the full ApoE gene. Fortunately these issues do not represent a weakness in the literature because such alleles are rare and predominant in non-Caucasian populations, which are on the contrary, the main “source” of the data.
even have started. As described above by the models of Lahiri (2007) and Jack et al. (2010), the first AD changes can be visible decades before the clinical onset. Based on this, effect of risk factors could be evident even at a young age, whenever these consist of the action of a genetic variable. By definition, the main function of a gene is to code for a protein and the expression of the ApoE gene results in the Apolipoprotein E polypeptide, whose main discovered function is to combine very-low/intermediate density lipoproteins that can bind to low-density protein receptors and allow the transport of lipids through the vascular system. The next paragraph will get more into detail of this molecular mechanism.

The human ApoE gene is located on chromosome 19 in position 3.2 (EntrezGene, updated on 8-Jan-2012) and is composed by about 3600 base pairs (including all uncoded exons) (Paik et al., 1985). The introns sections translate 299 amino acids and the abnormal version of the peptide coded from the $\varepsilon_4$ allele differs from the “classic” $\varepsilon_3$-derived polymer by one amino acid only, at positions 112 (arginine instead of cysteine). The sequence of base pairs (and therefore the sequence of amino acids) is directly responsible only for the primary level of protein structure; while the second level is mostly describable as $\alpha$-helix; there is further degree of discrepancy between ApoE $\varepsilon_3$- and $\varepsilon_4$-transcribed proteins in the third hierarchic level of structure. Most remarkably, the arginine-cysteine substitution causes a rearrangement of the side chain of arginine 61, which is more prone to interact with the amino acid at position 255 (Mahley & Rall, 2000). As a result, the main difference between $\varepsilon_3$ and $\varepsilon_4$ consists in the resulting position in space of the C-terminus, which is estimated to be located much more adjacent to the N-terminus in the ApoE $\varepsilon_4$ protein compared to the ApoE $\varepsilon_3$ protein, and confers different stability to the peptide (Saito, Lund-Katz, & Phillips, 2004). Since the residues in the C-terminus domain are those with major affinity to lipids and lipoproteins, there are mechanical differences in binding properties (and, therefore, in global functionality) between the two peptides (Ruiz et al., 2005). The ApoE displays a various set of functions in the human body, and these can be subdivided into two umbrella categories: lipid transport and neurobiological mechanisms (Mahley & Rall, 2000). The protein can bind to different classes of molecules, among which lipoprotein particles represent an important subtype. Different peptidic polymorphisms display different binding properties to lipoproteins: The ApoE $\varepsilon_3$-protein prefers to combine to High-Density Lipoproteins (HDL), whereas the ApoE $\varepsilon_4$-protein binds preferentially to Low-Density (LDL) and Very-Low-Density Lipoproteins (VLDL) (Hatters, Peters-Libeu, & Weisgraber, 2006). A lipoprotein is a molecule formed by a lipid kernel, surrounded by a surface made of other lipids, phospholipids and proteins, and it has an important role in lipid transportation, inflammation, general defence mechanisms, and binding to other proteins which, in turn, elicit their functions (Hoofnagle & Heinecke, 2009). Apart
from the well-known combination of HDL/LDL-VLDL to cholesterol in the determination of "good"/"bad" cholesterol, little is known about the functional differences of each class of lipoproteins. Unfortunately, little is known also about the mechanisms by which the $\epsilon_4$ allele increases the odds of developing AD, but it is likely that these neurobiological processes involve the main AD biomarkers (Mahley, Weisgraber, & Huang, 2006), and will be briefly mentioned later. Many are the interactions of ApoE with biological variables (Roses, 1997), and most of the evidence of its importance comes from animal research or studies in vitro, which lie outside the scope this descriptive context. There are some authors who, reviewing the evidence collected against ApoE, conclude that this gene/protein must be a key variable rather just a risk factor, for developing AD pathology (Crutcher, 2004; Raber, Huang, & Ashford, 2004), and it is directly involved in accentuating progressive failure in neuroplasticity (Teter, 2004).

1.4.2 APOE GENOTYPE AND IMAGING GENETICS

The branch of research named “Imaging Genetics” aims to investigate the effect of genetic variability on cognitive measures and susceptibility to disease (Mattay, Goldberg, Sambataro, & Weinberger, 2008). The starting point of this model is a trajectory of normal ageing defined in terms of structural, functional and behavioural alterations, moulded by inter-individual variability and other environmental (e.g.: lifestyle, diet, engagement in cognitive and physical activities…) and demographic (e.g.: gender, education, IQ…) factors. The theoretical leap from genetic makeup and cognitive discrepancies is quite big. A gene may have an unexpected impact on cognition, because “the closer we place our measurement to the level of the neuronal circuitry, the larger the effect of a single gene may be” (de Geus, Goldberg, Boomsma, & Posthuma, 2008), and cognition is not a direct effect of gene expression (Figure 1.2). As said before, the only function of a gene is by definition to bind to and code for its RNA counterpart and a protein, in order. Genes have an intricate impact on the normal trajectory of ageing: there are situations of abnormal ageing in which one single gene is responsible for the different ageing course (see for example Presenilin 1-2 and AD), but most times a single gene does not have an exclusive and well-defined impact on the investigated variables; Deary et al. (2004) hypothesised that different ageing pathways are the result of a multi-/oligo-genic influence. As a consequence, the impact of the single ApoE gene would tend to be estimated as not significantly relevant for AD onset (and for this reason it is often referred to as a simple risk factor).
Fig. 1.2 The principle of Imaging Genetics: different allelic makeups exert a cascade effect on many variables at different stages. The closer to gene expression the measurement point, the more precise and systematic is the effect on the variable of focus. From Callicott & Weinberger (2003)\(^5\) and Mattay et al. (2008)\(^6\).

The Imaging Genetics model warns against the variability by which the ApoE genotype is associated to cognitive skills, and predicts that the impact of the different alleles would be more accurately measurable if one investigated stages located theoretically “closer to the point of genetic expression”. For this reason, it is important to focus first on the relation between genotype and AD main biomarkers, in order to fathom whether first there is any specific mechanism that connects ApoE gene/protein to the peptidic markers, and second whether there is any significant statistical trend due to the possession of the $\varepsilon_4$ allele, rather than the “normal” configuration $\varepsilon_3\varepsilon_3$ for the same biomarkers and for the other important measures so far described, such as brain structure and brain function. It is also possible that

---

\(^5\) Callicott, J. H. and D. R. Weinberger (2003) "Brain imaging as an approach to phenotype characterization for genetic studies of schizophrenia" Methods in Molecular Medicine 77: 227-247. Figure 1: with kind permission from Springer Science and Business Media.

the ApoE gene displays a "butterfly effect", as suggested by Van der Flier (2012): according to her view the bigger the theoretical distance between the ApoE and the measured variable, the bigger its impact. This alternative idea however, does not argue against a big variability in the effect of the ApoE on distant variables such as cognitive skills.

1.4.2.1 APOE AND AD-RELATED PROTEINS

Aβ and TAU have been found to interact with the ApoE protein and whereas both human ε₃⁻ and ε₄⁻dependent lipid-free isoforms were reported to bind in vitro to Aβ₄₂ in the same residues that are responsible for binding to lipoproteins (244-272), the oxidated ε₄ isoform combines more rapidly than ε₃ (5 minutes rather than 2 hours) (Strittmatter et al., 1993). As parallel evidence, in another study NFT from human brain tissue were immunostained with dyes created to detect the presence of the ApoE protein. Staining was evident in proximity of the N-terminus (up to amino acid number 271, out of 299), but was absent for the C-terminus domain (last 28 residues), suggesting that the ApoE interacts with TAU in the formation of NFT or NFT-like filaments, and ε₄-protein is more susceptible to this process (Huang et al., 2001).

Based on this in vitro evidence highlighting the effect of different protein isoforms on the two main biomarkers of AD, it is expected that cross-sectional differences between carriers and non-carriers in the amount of central and peripheral deposition of these molecules in vivo and post-mortem should be reported.

No correlation was found between the presence of ε₄ and CSF Aβ₄₂ or t/p-TAU levels in a cohort of AD patients with an MMSE≈20 (van der Flier, Schoonenboom, Pijnenburg, Fox, & Scheltens, 2006). In a sample covering the whole spectrum of severity, significant reductions in CSF Aβ₄₂ concentration were found in carriers compared to non-carriers, without any t-TAU differences (Smach et al., 2008). A test-retest study did not find any effect of the allele on Aβ₄₂ of moderate AD in the long term, but found genotype-driven differences in t-TAU levels, with a steeper concentration increase in carriers (Kanai et al., 1999). Similarly, a cross-sectional association was found between ε₄ and CSF t-TAU in a bigger sample of AD of comparable severity (Tapiola et al., 1998).

CSF p-TAU was reported to be higher in carriers in a sample of MCI who later converted to AD-dementia (Buerger et al., 2005). Comparing baseline measurements of future converters and non-converters, the converters exhibited a significant relation between ε₄ and Aβ₄₂ but not between ε₄ and p/t-TAU (Herukka et al., 2007); in this study, no relationships were found in stable MCI. In another study, the MCI most likely to convert were ε₄ homozygotes with elevated p/t-TAU (Blom et al., 2009). The predictive power of genotype and CSF molecules on conversion was studied by Kester et al. (2011). Baseline levels of TAU and Aβ₄₂ and presence of
the $\varepsilon_4$ allele were significant predictors, as expected, but a model including genetic and cerebrospinal factors together revealed that $\varepsilon_4$ and TAU exert an independent effect, while $\varepsilon_4$ and $A\beta_{42}$ interact: $A\beta_{42}$ is a significant predictor in noncarriers only, and the $\varepsilon_4$ allele is a significant predictor only in case of normal baseline concentration of $A\beta_{42}$.

Differences were found between healthy carriers and non-carriers (age 21-88) for none of the variables (Vuletic et al., 2008) and for all variables in a sample including exclusively old participants, in the expected direction (Herukka et al., 2007), suggesting a relevant role played by age (neglected in the former studies). Only $A\beta$ and not p/t-TAU levels depended on the $\varepsilon_4$ presence in a third study (Liang et al., 2010). Age and genotype were found to be significant independent predictors of CSF levels of $A\beta_{42}$ only, without any effect on t-TAU, but only homozygotes were reported to differ from controls (Sunderland et al., 2004). Age was used as a variable also by Glodzik-Sobanska and colleagues (2009) in a regression model to explore possible interactions with the presence of the $\varepsilon_4$ allele in a cohort of healthy adults: whereas $\varepsilon_4$ was a significant predictor for t-TAU levels, both variables interacted to account significantly for variability observed in p-TAU; no effect was reported for the $A\beta_{42/40}$ ratio.

Tosun and coworkers (2010) tested $A\beta_{42}$, p-TAU and t-TAU in three samples (controls, MCI and AD) controlling for age, and found that the presence of the $\varepsilon_4$ allele correlated with $A\beta_{42}$ in all three subgroups and with both values of TAU in MCI only. Exactly the same pattern of results was reported by Blom and colleagues (2009), with an extra association between $\varepsilon_4$ and p/tTAU in controls.

In a study on healthy middle-age-to-old participants that controlled for the presence of the $\varepsilon_2$ allele, the ageing process and the $\varepsilon_4$ allele relationship was studied on both peripheral and central levels biomarkers: $\varepsilon_4$ was found to interact in an allele-number-dependent manner to predict CSF $A\beta_{42}$, but not p/t-TAU levels (Morris et al., 2010). Furthermore, these authors reported a global (ROIs were averaged) increased PiB binding rate for carriers as a function of age. Reiman et al. (2009) used both ROI and voxel-based approaches on their PiB data, finding that healthy $\varepsilon_3\varepsilon_4$ heterozygotes differed in tracer distribution from non-carriers, and this was even more pronounced in $\varepsilon_4$ homozygotes. Small et al. (2009) used $^{18}$F-FDDNP and maintained separate the ROIs to show that MCI diagnosis is linked to allelic differences in mediotemporal binding (greater retention in $\varepsilon_4$ carriers), whereas this trend was shifted in the frontal ROI for the healthy adults. A PiB tracer was employed by Drzezga et al. (2009) in AD individuals with an MMSE$\approx$20. All patients displayed the classic picture of temporoparietal, precuneal and PCC staining, but carriers had an extra involvement of temporoparietal and frontal voxels.
As neuropathological hallmarks of AD can generally be determined only with histological techniques, the effect of the $\varepsilon_4$ allele must be investigated in the role played on central deposition of AD-related neuropeptides. When the amount of AD-associated lesions was converted into Braak stages and plotted against age of death in a sample of histologically analysed brains, both NP and NFT stages were significantly associated with a younger age of death in carriers, suggesting that age is an important variable in the comparison between carriers and non-carriers (Ohm, Scharnagl, Marz, & Bohl, 1999). In a different design, histological assay of hippocampal and frontal parenchymal samples revealed no effect of allelic makeup on either NP or NFT, in samples of full-blown AD and controls (Landen, Thorsell, Wallin, & Blennow, 1996). In another study, $\varepsilon_4$ was reported to be related to NP and NFT deposition in frontal, temporal and fusiform cortices of brains affected by AD (Beffert & Poirier, 1996). Braak NFT stages were identified in a randomised sample of brains belonging to young individuals with absence of any NP involvement: the $\varepsilon_3\varepsilon_4$ genotype was positively associated with membership to this subgroup rather than that of brains with no NFT burden (Ghebremedhin, Schultz, Braak, & Braak, 1998). This finding on a group of young individuals (clearly years away from any early clinical manifestation of a form of sporadic AD) has led to the speculation that the $\varepsilon_4$ allele is causally linked to NFT deposition. The idea that the $\varepsilon_4$ isoform is actually an active variable that directly influences cerebral load of AD-linked proteins is supported by other findings. Brains of $>$60 year-old adults, not having any cognitive dysfunction at time of death, displayed allele-driven differences in NP deposition in frontal, temporal, hippocampal and entorhinal regions, but no NFT differences in any areas (Caselli et al., 2010). The lack of association between NFT and genotype was supported by another study (Kok et al., 2009), which also confirmed a difference in the overall cortical deposition of NP between carriers and non-carriers in all age subgroups, apart from the oldest (age: $>$90) and the youngest (age: 0-49) ones. In summary, it is difficult to extrapolate any evidence connecting NP and NFT deposition to the ApoE gene. One possible source of variability may be due to methodological differences, since some teams of authors preferred to look at the whole brain, while others focused only on specific brain regions.

It is vitally important to understand the impact of the $\varepsilon_4$ allele on the pathological hallmarks of AD, because the links allele-diagnosis and allele-neuropsychological profile appear to be no longer significant after controlling for NP and NFT lesions (Bennett et al., 2003). Splitting the carriers group into heterozygotes and homozygotes could help clarify this issue. Tiraboschi et al. (2004) performed autopsies on AD deceased individuals and did not find any NFT or NP differences between carriers and non-carriers, apart from a higher density of plaques in the superior temporal brain slice. When they further divided carriers according to the number of $\varepsilon_4$
alleles, they reported significant increases in $\varepsilon_4\varepsilon_4$ brains for NFT and NP in almost all sampled ROIs, without any proof of difference in cholinergic dysfunction.

The link between the ApoE and the cholinergic system is worth mentioning because the only approved treatments for AD are based on the neurotransmitter’s deficits associated with the disease. The molecular mechanism that could link the ApoE protein and the cholinergic system has been hypothesised to tap the phospholipids of the neuronal membrane, on which the availability of acetylcholine precursors depends. As already mentioned, the $\varepsilon_4$ isoform is connected to a qualitatively different binding to and transport of cholesterol and phospholipids; this could account for or interact with a less functional cholinergic innervation throughout the brain. $\varepsilon_4$-dependent reductions were registered in tissue coming from AD-brains compared to controls and AD $\varepsilon_4$ non-carriers in the hippocampus and temporal cortex (Poirier et al., 1995). Allen and colleagues (1997) analysed histologically the hippocampus, and temporal and occipital cortices of healthy and demented individuals. All areas were more severely affected in carriers than non-carriers, when compared to controls. Moreover, on comparing the presence and absence of $\varepsilon_4$ in the AD-only subgroup, just the hippocampus reached significance; a similar, but not significant trend was detected in the control-only subgroup. Other areas of tissue were selected by another team who found an effect of the $\varepsilon_4$ on frontal cholinergic activity (measured radiochemically), with higher severity in the $\varepsilon_4\varepsilon_4$ tissue (Soininen et al., 1995).

In summary, the ApoE $\varepsilon_4$ isoform appears to interact with all hallmarks associated with AD, albeit the specific molecular mechanisms are far from completely clear. Because of this, and because of the role played by other known or yet undiscovered variables, it is not possible to draw a main conclusion on the impact of the ApoE genotype on central and peripheral levels of $\text{A}\beta$, TAU and health of the cholinergic system.

1.4.2.2 APOE AND BRAIN STRUCTURE

There is a consistent body of research that has investigated the effect of the ApoE genotype on brain anatomy of older healthy and diseased individuals. Mostly, the interest has been on brain volume comprehensive of GM and WM, cleaned out of CSF; in this way, unless stated, researchers have been able to identify the role played by cerebral structures taken in entirety, again, voxel-by-voxel in an exploratory fashion, or ROI-wise.

Cross-sectionally, VBM analysis revealed that in moderate AD there seems to be no association between the $\varepsilon_4$ presence and general GM atrophy (Drzezga et al., 2009). The idea of the specific effect of the $\varepsilon_4$ allele presence on particular brain regions was tested in a sample of
moderate AD patients, in which the number of $\varepsilon_4$ alleles was a significant predictor for lower rates of cerebral volume in the temporal cortex and mediotemporal lobe, without any impact on the frontal lobe (Geroldi et al., 1999). The pattern of $\varepsilon_4$-linked mediotemporal atrophy was confirmed also in a slightly more severe AD sample (Barber et al., 1999). In a paper that excluded the $\varepsilon_2$ isoform, demented AD with an MMSE=19 strangely showed positive dose-dependent effect of the $\varepsilon_4$ on total brain GM volume (Yasuda et al., 1998). In a second study without $\varepsilon_2$ carriers, the same team (Hashimoto et al., 2001) scanned a larger cohort of moderate AD participants and calculated correlations between number of $\varepsilon_4$ copies and regional GM volume of key structures: a negative Spearmann $r$ was found for both hippocampal and amygdalar volumes in association with the allelic dose; intriguingly the results of the previous study were replicated for the whole-brain volume. A third study by this Japanese team (Mori et al., 2002) followed up a sample of AD for one year and reported cross-sectional effect of the $\varepsilon_4$ presence on hippocampal volume at any of the two measurements, but a significant impact on the rate of atrophic increase in this period in the expected direction; remarkably, no main effect of baseline AD severity and no genotype-severity interaction were significant. Finally, Pievani et al. (2009) found that their moderately affected AD sample displayed an interesting allele-driven rate of GM differences: $\varepsilon_4$ carriers had a reduced volume in medial and lateral temporal lobe bilaterally and in right occipital pole; in addition, they also displayed an increased volume in the PCC, the left lateral orbitofrontal cortex and the right dorsolateral frontal cortex. The emerging pattern describes an apparent role of the $\varepsilon_4$ allele on mediotemporal structures' volume, with sporadic evidence of involvement of those brain areas that have been normally connected to AD pathology, areas that could be particularly susceptible to the indirect effect of the ApoE genotype. The hippocampus seems to be a structure constantly mentioned by various teams, but unfortunately there is a degree of disagreement in the literature. Jack et al. (1998) found no effect of genotype on total hippocampal volume in moderate AD; moreover, genotype and hippocampal volume concurred independently to predict group (AD) membership. Basso et al. (2006) reported no difference in hippocampal volume when controlling for MMSE score (average score=17) between carriers and non-carriers, detecting instead a decreased volume of amygdalar structure in $\varepsilon_4$ carriers. Dose-effect GM reductions were also registered in milder (MMSE=22) AD, specifically in the hippocampus, parahippocampal gyrus, amygdala, fusiform and orbitofrontal cortex bilaterally (Filippini, Rao, et al., 2009).

There is similar research on MCI: Fleisher and coworkers (2005) described hippocampal volumetric differences between $\varepsilon_4$ homozygotes and non-carriers, with the difference between heterozygotes and non-carriers being significant in women only. A broad sample of healthy
elderly participants and MCI were genotyped and scanned to determine the effect of membership to the \( \varepsilon_4 \) carriers group (individuals with \( \varepsilon_2 \) alleles were excluded). Anatomical differences were reported in the \( \varepsilon_3 \) homozygotes for hippocampal and amygdalar volume, with no whole-brain difference (den Heijer et al., 2002). Hippocampal volume was smaller in MCI carriers in another study also (Farlow et al., 2004), and amnestic MCI carriers were found to have reduced grey matter in frontal and pericentral areas bilaterally and in right temporal cortex compared to non-carriers (Venneri, McGeown, et al., 2011).

ApoE \( \varepsilon_4 \) was also reported to have specific effects on structural decline in proximity to MCI-to-AD conversion: \( \varepsilon_4 \) carriers displayed decreased volumes in right temporal, right hippocampal and left insular GM in the twelve months prior to estimated conversion and decreased volume bilaterally in the insula, hippocampus and temporoparietal neocortex plus right caudate in the twelve months after conversion; likewise, a temporal and insular post-conversion volumetric loss was described for stable MCI carrying the \( \varepsilon_4 \) allele (Spampinato, Rumboldt, Hosker, Mintzer, & Neuroimaging, 2011).

Apparently, when scanning healthy people, results are different: a large sample selected randomly from a community of middle-aged adults was scanned and the outcome revealed no effect of the \( \varepsilon_4 \) allele on total, hippocampal and amygdalar GM (Cherbuin et al., 2008). Similarly, Jack et al. (1998) reported no difference due to genotype in their subgroup of healthy controls. Lemaitre et al. (2005) found no GM or WM total volumetric differences but reported a volume difference in both hippocampi between \( \varepsilon_4 \) homozygous and non-carriers, with no effect of one single allelic copy. Voxel-by-voxel analyses extended these findings with an additional significance reached in parahippocampal gyri and amygdalae. As contrary evidence, in a smaller sample of \( \varepsilon_2 \)-free \( \varepsilon_4 \) carriers and non-carriers, non-carriers had instead higher right total hippocampal volumes than carriers (Lind et al., 2006), replicating the findings of an older study (Tohgi et al., 1997). Right mediotemporal reduction, together with cerebellar, temporal and prefrontal GM clusters, analysed voxel-by-voxel, also discriminated carriers and non-carriers of a large sample comprising young and old adults without any \( \varepsilon_2 \) allele (Wishart et al., 2006). A recent study found global differences between Icelandic carriers and non-carriers in which carriers displayed smaller GM but normal WM levels, although they had larger amount of WM hyperintensities (Hafsteinsdottir et al., 2012).

In attempt to identify the pure effect of the genetic makeup, sets of monozygotic twins were analysed anatomically by Plassman and colleagues (1997). The group of carriers displayed significantly smaller hippocampi than non-carriers, suggesting an impact of genetics on structural maturation. Another twin (both monozygotic and dizygotic) study confirmed the
degree of ApoE-dependent heritability of structural measures, this time computing cortical thickness (Kremen et al., 2008).

Cortical thickness has been studied by other teams in relation to the presence of the $\varepsilon_4$ allele. Carriers were described as having thicker GM cortex in occipital, temporal and frontal regions bilaterally, and in the right parahippocampal gyrus; vice versa, there was a steeper thinning in carriers’ similar areas as age increased (Espeseth et al., 2008); no total or WM/GM volumetric differences were reported, though. More detailed descriptions of the involved structures were published by Burggren and colleagues (2008) who scanned old healthy $\varepsilon_3$ homozygotes and $\varepsilon_3\varepsilon_4$ heterozygotes using a ROI approach, computing a continuous strip of hippocampal GM and targeting each subfield. In this more detailed exploratory design, whereas volumes did not display any difference, thinner values associated with the $\varepsilon_3\varepsilon_4$ genotype were found for the subiculum and the entorhinal cortex.

This finding was confirmed in a longitudinal study, in which healthy adult female $\varepsilon_4$-carriers were reported to have lost a bigger amount of hippocampal volume than non-carriers after two years, but there was no difference in total cerebral volume (GM + WM) change between groups (Cohen, Small, Lalonde, Friz, & Sunderland, 2001). A TBM approach aimed at determining the rates of increase of atrophy after 5 years from baseline in a sample of healthy heterozygotes only ($\varepsilon_2\varepsilon_3$ and $\varepsilon_3\varepsilon_4$) showed that only the temporal lobes and the right hippocampus exhibited differences indicating decline (Lu et al., 2011). In another study, no volume changes were found to be related to the $\varepsilon_4$ allele in the hippocampus of controls and MCI after one year, but a comparison between $\varepsilon_3\varepsilon_3$ and $\varepsilon_3\varepsilon_4$ individuals with a diagnosis of AD revealed a steeper one-year decrease in the latter group (Schuff et al., 2009). Donix et al. (2010) segmented hippocampal WM and CSF out and studied volume changes after two years. They found diminished thickness in the carriers’ CA1, subiculum, entorhinal cortex, fusiform, parahippocampal and perirhinal cortices, and total hippocampus, with no changes in non-carriers. Two-to-three years elapsed from baseline to the follow-up scan in another study (Moffat, Szekely, Zonderman, Kabani, & Resnick, 2000). Apart from finding no differences at baseline, the authors described an $\varepsilon_4$ status-dependent annual amount of hippocampal volume loss, with carriers displaying a faster rate. No differences were reported for total brain volume and hemispheric preference though. A far larger cohort of healthy elderly persons were scanned at baseline and after four years, revealing hippocampal reduction in both heterozygotes and homozygotes for the $\varepsilon_4$ allele, and a total-brain GM reduction in homozygotes only (Crivello et al., 2010).
CSF values of Aβ_{42} and brain atrophy were co-measured in a sample of MCI, and an independent impact on the ε_{4} allele and Aβ_{42} was registered, with no significant interaction (Tosun et al., 2010).

The specific influence of genotype was also studied in relation to WM and WM changes with MRI and DTI-MRI. Just a few of these studies will be mentioned here to clarify the role of the ε_{4} allele in relation to the WM of healthy middle-age-to-old adults. Using standard structural MRI, Ready et al. (2011) found smaller WM total volume in carriers. Persson and colleagues (2006) explored the effect of the allele voxel-by-voxel and found decreased anisotropy in posterior sections of the corpus callosum, in the occipito-frontal fascicula and in the left hippocampus. In a similar study, age differences were found in which older carriers exhibited MD increases compared to non-carriers, and young carriers displayed FA decreases compared to non-carriers (Heise, Filippini, Ebmeier, & Mackay, 2011); both increases were widespread. Ryan et al. (2011) reported an effect of the ApoE ε_{4} allele on diffusion rates throughout the whole selected ROIs, but the main effect on anisotropy was limited to frontal WM and to the splenium of the corpus callosum. Moreover, the ε_{4} allele played a modulatory role on the normal ageing effect exhibited on anisotropy in temporal and frontal WM. Finally, a single ROI was traced by Wang et al. (2012) in the parahippocampal WM of mild AD and controls altogether. The ε_{4} allele was a significant predictor for WM volume and MD but not for FA. The involvement of the corpus callosum is in line with another study in which this structure was parcellated according to the cortical region the WM fibres were directed to: only the "prefrontal" portion of the corpus callosum was affected by the genotype, with ε_{4} carriers displaying a steeper age-related shrinkage than noncarriers (Filippini, Zarei, et al., 2009).

CDR=0 and ε_{2}-free participants over 60 years old were tested to compare the effect of ApoE on GM in relation to WM structure (Honea, Vidoni, Harsha, & Burns, 2009). Carriers displayed smaller volume in the left lingual, parahippocampal and angular gyri, left superior frontal cortex and right inferior parietal cortex and precuneus. They also displayed bigger volume in right frontal areas, hippocampus and precuneus, and in left middle temporal and occipital gyri; in the same sample, DTI revealed just a drop in anisotropy in the carriers’ left parahippocampal gyrus.

These studies display an overall variable pattern of findings. Differences in results might be due to the confounding impact of factors such as age, gender, severity of pathology, and lack of statistical power due to small sample sizes (Hashimoto et al., 2001). Another reason is the heterogeneity of the MCI samples, since often stable MCI are mixed with progressing MCI. Moreover, a control for the third isoform of the ApoE gene is rarely carried out: the impact of
the $\varepsilon_2$ allele might be consistent both in the control ($\varepsilon_2\varepsilon_2$ and $\varepsilon_2\varepsilon_3$ genotypes) and the experimental group ($\varepsilon_2\varepsilon_4$ genotype). Any form of control for this variable has been consistently pinpointed in the text for this reason. Another aspect to take into account is the possible presence of type II errors in this kind of studies. When associated with another variable, the volume of a small structure such as the hippocampus can display only very small differences. It is therefore possible that a cross-sectional study is not able to capture these minute differences, especially if the group is heterogeneous (Mori et al., 2002). The same can be extended to subtle cognitive differences that lie on the thin line between absence of deficit and dementia. This may be a probable reason why no effect of the genotype on cognition has been reported in studies in which the focus was on parameters of brain structure.

Structural differences related to genetics have rarely been studied at young age and in those rare cases volumetric differences in GM and WM volumes have been either found (Shaw et al., 2007) or not found (Heise et al., 2011). It can therefore be speculated that measures of cerebral structure in our possession today are not sensitive enough to detect systematic $\varepsilon_4$-dependent differences at young age. Functional neuroimaging represents a frontier of research that could be more sensitive to detecting genetic differences in early adulthood. In addition, testing young adults would also allow one to understand what the effect of the genetic makeup is on pure brain function, without any modulation played by important structural damage or cerebral deposition of neurotoxic molecules, visible instead at a more advanced age.

1.4.2.3 APOE AND BRAIN FUNCTION

The three main techniques of functional neuroimaging (PET, SPECT, fMRI) have been studied in relation to the ApoE genotype and will be reviewed here; the body of research using PET and SPECT techniques has been limited compared to fMRI.

To date, many papers have investigated the association between allelic configuration for the ApoE gene and brain function during a cognitive task. Unfortunately results show ample inconsistencies: in a recent paper, Trachtenberg and colleagues (2012) literally struggled in the attempt to extrapolate a well-defined conclusion on the effect of the $\varepsilon_4$ allele on neural activity following a task computation. The main problem appears to be the lack of a systematic paradigm. Too many tasks have been employed in this kind of research, and too much variability has prevented researchers from reaching unambiguous conclusions. Those authors however reported that there is some sort of undetermined difference in the working brain carrying the $\varepsilon_4$ allele. Recent studies found task-related functional differences related to age: In two sequential papers, Filippini et al. (2011; 2009) used the same encoding task to fathom
the impact of the \( \varepsilon_4 \) isoform on neural function at different points over the age span: whereas young carriers displayed increased activation in mediotemporal areas compared to non-carriers, the same procedure elicited an opposite trend in older adults, suggesting a different role possibly played by the same variable at different steps of development. It is interesting to note how this genetic variable exerts its effect on neural function decades before any possible clinical onset of neurodegenerative disorders even in the absence of other markers.

Unfortunately, as pointed out by Trachtenberg and colleagues (2012), there is no task that has been identified as a gold standard, to keep the methodology constant. Fleisher et al. (2009) pointed out that task-associated fMRI scans depend highly on the physiological baseline level of the brain: ApoE \( \varepsilon_4 \) carriers therefore, having different baseline parameters, may display differences during task performance for this reason rather than other neural mechanisms linked to different ways to compute the mental processes associated with the task. As a normal physiological characteristic of the central nervous system, this baseline variability of brain function should be evident even in genetics-free fMRI studies; the insertion of this genetic variable, introduced for other background reasons (in this case its association to AD), would improve the model, but would intervene on the functional outcome for a different reason which is not of our interest.

In this scenario the most relevant source of information comes from functional studies that, paradoxically, did not use any task: as previously reviewed, PET and SPECT procedures in AD have been usually carried out at rest, unveiling a pattern of findings that is relatively consistent across studies. Some papers have extended this rationale incorporating a genetic variable. Sakamoto and colleagues (2003) investigated baseline and follow up global and regional brain perfusion of patients with MCI at baseline and AD at the 2-year follow up. What they found was that carriers had reduced rate of regional perfusion in the right parietal lobule at baseline and after two years in frontal areas compared to control patients. The frontal lobe was also identified as the core of differences in perfusion between moderate AD \( \varepsilon_4 \) carriers and non-carriers also in another study, in which the authors conclude that the frontal lobe could be the locus of ApoE allelic differences in AD (Hogh et al., 2001).

Switching to PET, a study of a sample of mild AD revealed that \( \varepsilon_4 \) carriers had lower metabolism in right ACC and PCC and in bilateral frontal areas. Moreover, the presence of the allele exacerbated the effect of ageing on the ACC and medial frontal metabolism (Mosconi, Sorbi, et al., 2004). The same team also tackled conversion from MCI to AD with regards to ApoE genotype: a one-year follow-up spotted all the post-pre differences in metabolism, and the impact of the \( \varepsilon_4 \) allele on this differential measure was linked to metabolism in the left ACC and inferior frontal gyrus, which was lower for carriers (Mosconi, Perani, et al., 2004).
Furthermore, there were genetic differences even in the group of non-converters taken on its own. The regional rate of metabolism was reduced in carriers belonging to this set of (at that time) un converting MCI in the middle temporal gyrus, inferior parietal lobe and precuneus bilaterally. In another study, a sample of non-demented adults was scanned at baseline and after two years: $\varepsilon_4$-associated lower metabolism rates were recorded in the left inferior parietal lobe, lateral temporal cortex and PCC using a voxel-based processing. After 2 years a decline was reported in the frontal lobe for non-carriers only, and in inferior parietal/lateral temporal cortices for carriers exclusively (Small et al., 2000). Contrary to the studies previously mentioned in which the frontal lobe was the positive kernel of genotype-due differences in AD, in this paper the impact of the $\varepsilon_4$ allele on frontal metabolism appears to be reversed, with non-carriers displaying increased decline. This evidence does not find support in the findings of another research team. In a set of studies, Reiman and his team used PET to explore the rate of cerebral metabolism of healthy adults with a family history of AD: $\varepsilon_4$ carriers displayed no change in global brain metabolic rate, and instead had a pattern of reduction in the PCC, prefrontal, temporal and parietal areas, similar to that observed in a group of probable AD, with additional prefrontal clusters (Reiman et al., 1996). A 2-year follow-up revealed that non-carriers continued declining in ACC, PCC, parietal cortex, and caudate nucleus. This drop was significantly steeper for carriers in all these areas, plus an involvement of prefrontal cortex, parahippocampal and lingual gyrus, thalamus and basal forebrain (Reiman et al., 2001). From both studies it emerges once again how prefrontal areas are particularly sensitive to the presence of the $\varepsilon_4$ allele in terms of functional decline. Twenty-to-thirty-nine years-old carriers and non-carriers were then compared to investigate the effect of ApoE, leaving the process of ageing out: once again no effect of $\varepsilon_4$ was found on total rate of metabolic decline, and once again carriers displayed reduction in FGD regional uptake in parietal, temporal and prefrontal areas together with the PCC (Reiman et al., 2004). In the same paper, the conjunction of the young and the previously published old sample confirmed these results across the whole life timeline, boosting the evidence that the biggest difference is ascribable to prefrontal areas and ACC. These results were eventually expanded with evidence of a dose-effect of the $\varepsilon_4$ allele (Reiman et al., 2005). Apparently this pattern of variable resting-state cerebral metabolism is not culture-bound, as the same team (working in the USA) replicated their results in a community of Latino Mexican Americans. No ethnicity-by-genotype interaction was found but there was a main effect of the $\varepsilon_4$ presence only on regional metabolic rate in the PCC, precuneus, parietal cortex, ACC, hippocampus and thalamus (Langbaum et al., 2010).

A further study on young adults used $H_2^{15}$O PET to estimate cerebral blood flow: findings revealed decreased regional perfusion in bilateral temporal gyrus but also increases in left
insula, right supramarginal gyrus and inferior occipital gyri (Scarmeas, Habeck, Stern, & Anderson, 2003), a type of findings unreported by other studies. This sample was however relatively small, as 14 non-carriers were compared to 3 carriers only. H$_2^{15}$O PET was also used by Thambisetty and colleagues (2010) who followed a sample of older adults up for about 8 years. They detected a decreased rate of regional blood flow in left frontal areas, right orbitofrontal cortex, bilateral temporal areas and the right superior parietal cortex in carriers over time, with a parallel blood flow increase in the right insula.

The principal functional neuroimaging technique used nowadays is fMRI. Similarly to PET and SPECT task-less paradigms, fMRI has been recently used to measure functional parameters of the brain at rest and to detect ApoE-driven differences. Persson and colleagues (2008) scanned a sample of adult and old participants while performing a semantic memory task. They chose to investigate the task-related deactivations (subtracting the rest scan from the task scan) and found weaker activity for carriers in medial prefrontal cortex, left mediotemporal cortex and right parietal ROIs, findings confirmed by voxel-based exploratory analyses. Pure measures of brain activity at rest are preferred to deactivations, though, and very recent publications have used this methodology, reporting that the $\varepsilon_4$ allele does not seem to affect DMN measured in terms of functional connectivity at old or adult age (Koch et al., 2012; Trachtenberg, Filippini, Ebmeier, et al., 2012). Another team has reported instead a decreased connectivity for 50-to-60 year old carriers in the precuneus and rectus gyrus, with a parallel increase in correlational activity between temporal and prefrontal clusters (Fleisher et al., 2009). In this study, however, all carriers also had a family history of AD, and it is not clear how much variability can be accounted for by this status. Correlational maps revealed an effect of the $\varepsilon_4$ presence on precuneal connectivity in another study as well, which specifically tested adults with low PIB retention and high CSF A$\beta$ concentration. Increased connectivity between precuneus and occipital areas was observed for carriers, while decreased connectivity was evident from precuneus to middle temporal cortex, ACC, left hippocampus and left parahippocampal gyrus (Sheline, Morris, et al., 2010).

To my knowledge no study has specifically targeted regional homogeneity in relation to the ApoE, whereas there is some evidence emerging from a somewhat similar but different statistical technique. Two teams carried out a dual regression analysis, a procedure that aims to elaborate a pattern of connectivity for every single voxel. A group-level map of components is then estimated to fit each participant’s resting-state BOLD signal; this map is interpreted as a picture of how synchronous the BOLD signal is in a group, and two groups can be statistically compared in terms of this map. Westlye et al. (2011) scanned $\varepsilon_4$ carriers and non-carriers in their sixties, and reported different patterns of regional synchronisation; carriers displayed
augmented synchrony in BOLD signal within regions that are part of the DMN. The authors interpreted these findings as a failure of standard processes of neural decoupling. Normally, an atom and its electric charges exert an influence on neighbouring atoms as a response to the magnetic field; this translates into a synergic response in terms of a BOLD signal that is not homogeneous anymore but becomes converted into multiple peaks (coupling). “Decoupling” indicates the processes that eliminate these artifactual peaks and smooth the signal. A problem in decoupling would result into an increased synchronisation of BOLD signal and would suggest a neural difficulty. It is not clear to what extent reduced functional connectivity (and a reasonably expected reduced regional homogeneity) can conciliate from a theoretical point of view with a value of increased synchronisation observed in carriers, but the overall interpretations given by the authors suggest in all cases a neural impairment evident in the presence of the $\varepsilon_4$ isoform. A dual-regression-based methodology yielded the same pattern of results in a sample of young adults (Filippini, MacIntosh, et al., 2009). This study, together with the research assessing cerebral metabolism in young adults (Reiman et al., 2004), presents a strong pillar of evidence that suggest the presence of neural differences from a young age. The literature supports this idea bringing more evidence through functional studies with a task to perform (Scarmeas et al., 2005), and differences in entorhinal cortex thickness were also found in a sample of even younger participants (children and adolescents) (Shaw et al., 2007). It is likely that any variability found at college age or even at younger age is not accounted for by the classic features normally associated to AD, like structural changes or molecular hallmarks but rather reflects a pure expression of the ApoE genotype. One recent paper investigated the effect of genotype on functional connectivity of a big sample of healthy adults: carriers showed decreased PCC-seeded functional connectivity (used as a model for the DMN) at rest and increased ACC-seeded connectivity (used as a model of another network normally anti-correlated with the DMN), supporting the pattern of differences so far described (Machulda et al., 2011).

In summary, despite a few discordances, studies of functional neuroimaging display a consistent pattern of results: the presence of the $\varepsilon_4$ allele is associated with decreased brain function in areas normally belonging to the DMN, i.e. posterior neocortex, mediotemporal regions, PCC/precuneus, ACC and frontal clusters. At first sight, the idea of functional neuroimaging being more sensitive to genotype than brain deposition of NP and NFT seems to go against the prediction made by the Imaging Genetics model. The “dogma” of this model states that the farther the link between the gene and the variable of interest, the more variable the results (Mattay et al., 2008). Brain function and the ApoE gene are actually theoretically distant, but: 1) we do not know yet to what extent the ApoE protein affects Aβ
and TAU: they all work at a molecular level but this does not necessarily imply that they are causally related; 2) brain function can boast an accurate measurement paradigm and explanatory framework, which is the DMN. The mechanisms associated with location of AD-related lesions in the brain and in the CSF have not been completely discovered, and this lack of knowledge may undermine the “trustworthiness” of the ApoE-NP/NFT link, in favour of a more accurate ApoE-DMN link. In brief, functional neuroimaging appears to have more satisfying investigation methods; 3) the presence of alterations in connectivity between the precuneus and other brain structures in the absence of significant Aβ deposition (Sheline, Morris, et al., 2010) suggests an independent influence of these two variables on brain function. These points explain why accuracy of brain function and inaccuracy of NP/NFT measures do not clash with the Imaging Genetics model.

1.4.2.4 APOE AND COGNITIVE FUNCTION

Task-dependent functional neuroimaging does not represent the best paradigm to study the impact of the ApoE isoforms on neural function (Trachtenberg, Filippini, & Mackay, 2012) as there is uncontrollable source of variability from procedure to procedure and from task to task. Based on this, pure behavioural studies do not appear to be of primary relevance for defining the differences between Ɛ4 carriers and non-carriers at various stages of the pathology, as a huge degree of variability and inconsistencies is expected. In support of their Imaging Genetics model, Mattay et al. (2008) explained how cognitive measures are complex and can be influenced by variables such as the level of cooperation or the use of various strategies. This makes the scenario more complicated when trying to understand the impact of a single gene. Nevertheless, it is interesting at least to list the main findings in order to identify a possible rough trend that represents the relationship between the opposite edges of the Imaging Genetics model.

Observational studies pinpoint how the presence of the Ɛ4 allele increases the odds of conversion from MCI to dementia of AD type (Keage et al., 2010; Wang, Hong, Lin, Liu, & Chen, 2011). More detailed attention on cognitive domains has been mostly paid to samples of healthy adults, but these findings are very difficult to interpret. Overviewing this field of research, dozens of papers have been published which report analyses carried out to detect an effect of the ApoE genotype on cognitive skills, and a substantial part of the neuroimaging investigations previously reported included pure behavioural data as well. Most of the findings come from samples of old non-demented individuals. A good starting point to understand the effect of the gene on cognition is represented by meta-analyses: Small and colleagues first (2004), and Wisdom, Callahan & Rosnick more recently (2011) reviewed 38 and 77 older
papers with meta-analytical procedures. In general terms, the ApoE $\varepsilon_4$ allele seems to negatively influence measures of global cognition, episodic memory and executive functioning, with a modulating role of age and zygosity. There seem to be areas of cognition which are sensitive to the presence of the $\varepsilon_4$ isoform, but the picture looks more complicated. Meta-analyses depict just the general trend of a large sample of studies, therefore it is natural to find articles in agreement with the findings of those meta-analyses (Fleisher et al., 2005; Honea, Vidoni, et al., 2009; Westlye et al., 2011; Zehnder et al., 2009) and other papers that did not find any effect of genotype on cognitive functions (Bennett et al., 2003; Bunce, Fratiglioni, Small, Winblad, & Backman, 2004; Crivello et al., 2010; Jorn et al., 2007; Kim et al., 2002).

Similarly, longitudinal changes due to the $\varepsilon_4$ allele have been sometimes documented (Small et al., 2000), and sometimes not (Cohen et al., 2001). A recent observational paper did not find any longitudinal impact of ApoE on any of the measured variables after a 12-year follow-up, at any of the age ranges included in the sample (Van Gerven, Van Boxtel, Ausems, Bekers, & Jolles, 2012). The main point appears to be the relatively small effect the $\varepsilon_4$ allele would have on behavioural measures.

An alternative interpretation of the weak association between genotype and cognition could be the inappropriateness of tests and tasks normally used in clinical and experimental settings. It is possible that performances of non-demented adults display ceiling effects, and for this reason the true effect of the $\varepsilon_4$ allele is not detected. This idea is evidenced by the results reported by authors who used different tasks created to tap specific cognitive components. A version of the Stroop test based on simultaneous inhibition and switching skills time was chosen by Wetter and colleagues (2005) and carriers made significantly more mistakes than non-carriers. The use of this same task revealed a complex interaction between gender and genotype in a more recent study (Reinvang, Winjevoll, Rootwelt, & Espeseth, 2010). In this paper performance on the Letter Number Span Task from the WAIS (targeting various aspects of working memory at the same time) was also impacted by the presence of the $\varepsilon_4$ allele, but strangely heterozygotes had poorer performance than homozygotes. In other papers $\varepsilon_4$-carriers displayed significant lower scores on prospective memory measures (Driscoll, McDaniel, & Guynn, 2005), disengagement of spatial attention and working memory (Greenwood, Lambert, Sunderland, & Parasuraman, 2005), object recognition and spatial navigation (Berteau-Pavy, Park, & Raber, 2007), and on a word-recall task in which divided attention was necessary to solve simultaneous arithmetical equations (Rosen, Friz, Putnam, Harwell, & Sunderland, 2002). In addition, O’Hara et al. (2008) found no differences in the standard pencil-and-paper DSMT but a difference emerged in a computerised version of the same task which recorded Response Times (RT) to the millisecond. It is really hard to picture a
precise trend of results from these studies, but the emerging idea is that the effect of the ApoE genotype on cognition could be very subtle and you would need either particularly challenging tasks or more sophisticated cognitive instruments to detect it.

There is evidence suggesting that the effect of the $\varepsilon_4$ allele on cognition may be the result of the interaction between the allele and another variable. For example, histological evidence suggests that cognition at a later stage of life depends on the combined effect of the $\varepsilon_4$ presence and $A\beta_{42}$ load; the relation between the $\varepsilon_4$ allele and cognition was no longer significant after adjusting for $A\beta_{42}$ load (Bennett et al., 2005).

The ageing process is another important variable. When divided into different age groups, nondemented adults were differentially affected by genotype, with older-old adults being negatively influenced (Baxter, Caselli, Johnson, Reiman, & Osborne, 2003; Liu et al., 2010; Nilsson et al., 2006) when compared to younger elderly, and, inversely, with young adults being even positively affected: healthy carriers in their twenties/thirties were reported to have higher IQ (Schultz et al., 2008; Yu, Lin, Chen, Hong, & Tsai, 2000) and better delayed recall (Mondadori et al., 2007). These findings have been supported by neuroimaging findings which attested that young carriers have enhanced task-associated functional connectivity between the mediotemporal lobe and the PCC/limbic regions (Dennis et al., 2010) and presumed enhanced EEG-associated brain function during a working memory task (Alexander et al., 2007). The idea of this allele being beneficial at a young age and detrimental as the age increases is compatible with a model of antagonistic pleiotropy (Han & Bondi, 2008). This idea is intriguing but maybe too hasty, because Bunce et al. (2011) found no difference between young carriers and non-carriers in a large sample (N>5000) after the administration of a set of tests tapping aspects of attention and memory, nor did Ihle and colleagues (2012) in a recent meta-analysis.

The overview so far is puzzling, and semantic fluency tasks represent the best example of how confusing this literature can become. Healthy middle-aged-to-old $\varepsilon_4$ carriers produced significantly fewer words in the “animals” category (Miller, Rogers, Siddarth, & Small, 2005). Carriers with a family history of AD exhibited the same difference with three categories: animals, fruits and vegetables (Fleisher et al., 2009). Similarly, in a 10-minute category fluency task, carriers produced significantly (one-tailed $p$ value) fewer words and clusters of words than non-carriers (Rosen et al., 2005). Alexander et al. (2007) analysed their data obtained from various age-strata of the population and found opposite results. $\varepsilon_4$ carriers outperformed non-carriers in the semantic fluency task (and the difference was tracked post-hoc in the 51-65 y.o. subgroup). Furthermore, Thambisetty and colleagues (2010) reported that even
longitudinally, non-carriers had steeper decline than carriers in category fluency after an average of 7-8 years. Analysis of qualitative parameters of semantic fluency proved to be sensitive to the initial effect of AD on cognition (Forbes-McKay et al., 2005), but this was not performed in relation to the ApoE genotype in a sample of healthy adults. In a recent study this hypothesis was tested in a sample of MCI. Fewer words were produced by carriers, and a parameter of estimated familiarity appeared to differentiate carrier MCI from controls, but no difference was reported between controls and MCI with no $\epsilon_4$ (Biundo et al., 2011). Lower scores of cognitive tests were reported even among carriers having MCI (Farlow et al., 2004) or subjective memory complaints (Wehling, Lundervold, Standnes, Gjerstad, & Reinvang, 2007). In another study, regardless of conversion to dementia, the only task in which carriers had worse performance than non-carriers was the delayed recall of the Rey Figure (Mosconi, Perani, et al., 2004). Finally, baseline data of Whitehair et al. (2010) revealed how carriers had worse cognitive functioning in the ADAS-COG battery, DSMT, MSSE, digit cancellation, category fluency, a series of tests measuring verbal delayed recall and CDR sum of boxes, task that was instead unaffected by genotype in a sample of progressing MCI (Spampinato et al., 2011). Heun et al. (2010) recruited healthy participants and monitored the proportion of decline to MCI, finding no association with genotype after four years and a half.

The last group in which the effect of ApoE on cognition should be examined would be full-blown AD, but unfortunately by this stage a substantial portion of cognitive resources is no longer available, and each between-group comparison should be interpreted with caution. No differences in severity or duration were found between AD carriers and non-carriers (Landen et al., 1996) and no baseline cognitive differences were registered between carriers and non-carriers in the AD sample of Cosentino and co-workers (2008). Along the axis of time, there are authors that have reported no impact of the ApoE genotype on cognitive decline in AD (Growdon, Locascio, Corkin, GomezIsla, & Hyman, 1996), and there are samples of AD displaying differences in the overall cognitive decline according to presence or absence of $\epsilon_4$ alleles, with carriers having a steeper negative regression slope, interpretable as faster decline (Cosentino et al., 2008). Martins and colleagues (2005) published evidence that a non-linear model can fit the data better than regression lines. Trying to bypass this confusion, Davidson et al. (2010) carried out a Latent Class Analysis (similar to Principal Component Analysis) to identify the various cognitive scenarios through which AD presents itself in relation to the ApoE genotype. Carriers were more likely to have a uniform pattern of impairment across various cognitive domains rather than displaying difficulties in a limited set of skills like concentration/attention or memory.
As a final remark, no overall effect of genotype was noticed in a comprehensive sample of controls, MCI and AD with regard to a full neuropsychological battery (Bigler et al., 2003).

To summarise the branch of research investigating the impact of the ApoE makeup on cognitive abilities, the “prophecy” put forward by the Imaging Genetics model has come true: behavioural measures represent a stage in which too many variables concur to play their own role or in interaction with one another. From this point of view, the variability in the findings does not permit the detection of a strong pattern of conclusion in any of the progression stages of AD pathology (normality, MCI and dementia).

In conclusion, it is likely that different alleles for the ApoE gene display an effect on important variables associated with AD at various levels, from specimens to cognitive functions. However, since the true mechanisms by which genotype exerts its influence are not known at present, we are forced to treat it as a variable in association rather than in a relation of causality with the dependent variables that have been measured. Observational evidence does not permit the drawing of precise conclusions on the link between ApoE and Aβ/NP, TAU/NFT, brain structure or cognitive functions. The only pattern of findings that seems consistent is the disruption of DMN hubs associated with the presence of the Ɛ₄ allele.
Ageing consists of a multidimensional set of changes involving, among others, behavioural measures, brain structure and brain function (Deary et al., 2004). There has been interest in detecting any variables that could slow down the detrimental aspects of ageing, help protect from, or delay the onset of neurodegenerative processes, and/or slow down the associated decline. Some aspects of lifestyle are believed to be healthy and protective against the effects of ageing on brain and cognition. This belief has been in the first instance the result of simple common sense, but in recent times it has also received attention from the scientific community. A popular view hypothesises that at least two lifestyle aspects are specifically protective against the detrimental impact of MCI and AD. These two factors are the engagement in physical activity (PA) and in cognition-stimulating activities (CA).

### 2 PHYSICAL ACTIVITY

#### 2.1 INTRODUCTION

The effects of PA on brain and cognition represent an ample topic that deserves attention from a multidimensional perspective: exercise exerts an uncountable number of effects on human physiology, and the interest here is in those aspects that are related to normal and pathological ageing, with a focus on MCI and AD.

PA is a category that defines a large set of entities: examples of PA include walking, running, resistance training, strength training, and calisthenics. Some authors include leisure activities such as gardening or dancing in this category as well. It is difficult to draw a precise line between what can be considered as PA, and what simply involves movement. Most of the times, the purpose of promoting PA in an ageing population is enhancing their general well-being and independence from a holistic perspective. The effects of PA on human physiology are multiple, and most reviews highlight its multifaceted benefits on the ageing process. Autonomous nervous system, hormones, diet, smoking habits, social aspects, diabetes, heart disease, mobility, are some of the specific variables that have been heterogeneously studied in relation to PA, and most of the time general terms like “healthy ageing”, “perceived well-being” or “successful ageing” are referred to in an overview of all these parameters. Among
the multiple benefits, the role played by PA on neural and cognitive domain deserves particular attention. First of all it is necessary to classify PA into categories. The main distinction that can be universally made in the literature studying PA in relation to brain and cognition distinguishes two separate paradigms: Acute Exercise (AE) and Chronic Exercise (CE). AE aims to create a temporary perturbation in the neural system, in order to investigate its immediate effects. A short bout of exercise is administered, and neurocognitive variables are measured before, during, or after the session. Variability in this paradigm lies in the type, intensity and duration of the session, the time-point of measurement, and variables associated with the individual (e.g. fitness levels, age, gender ...). CE consists instead of a prolonged period of training, where variables are measured at its immediate end, and usually compared to a pre-training baseline (and ideally to a control group). AE usually has the experimental purpose of studying the temporary impact of a condition of effort on a normal system, in order to understand its immediate consequences. A within-subject design can be easily carried out, in which participants also perform a control condition (e.g. at rest); it is generally not time-consuming for the participant and it allows the researcher to recruit a large number of participants. CE is normally used to induce a long-lasting change in the neural system, which would translate into a benefit associated with enhanced neural and neuropsychological variables; the design for these studies is usually mixed and the number of participants is limited, as participation in this sort of research is time-consuming and a prolonged commitment is required of the participants. A parallel branch of research has investigated a more general lifestyle engagement in physically stimulating activities without manipulation of CE features. These are longitudinal studies in which the amount of PA is estimated in an observational way and treated as a predictor for one or more dependent variables; usually these studies are less precise because there is a clear lack of control on the PA variables, but a large number of participants is usually recruited, which boosts statistical power. In theory, AE and CE are expected to be qualitatively similar, as they technically differ only in terms of number of sessions, as CE is a sequence of AE. For this reason, it appears theoretically acceptable that the two paradigms may share part of the interpretational mechanisms.

Another important distinction often described in papers is that between aerobic and anaerobic exercise. Usually aerobic exercise comprises activities that involve a large consumption of oxygen such as running or cycling, whereas anaerobic exercise describes activities such as strength training that focus on short-term muscular work and do not lead to a condition in which a lot of breathing is required to recover from effort. This is a rough categorisation that does not exactly coincide with physiological evidence: the study of metabolism during exercise shows that there are three major pathways for the production of energy at a muscular level:
oxidative, phosphagen and glycolytic. Wells, Selvadurai, & Tein (2009) describe these three metabolic responses as different pathways to produce Adenosine TriPhosphate (ATP), a molecule often referred to as the "metabolic token" of human body; the breakdown of this molecule results in the loss of one phosphate group (the molecule becomes Adenosine DiPhosphate (ADP)) and it liberates energy that becomes available for work.

**Fig. 2.1** The timeline of metabolic response to exercise: Marathon running and cycling are often referred to as aerobic activities because, after five minutes, they exclusively activate an oxidative response in the body. The other two pathways are considered anaerobic because they do not use oxygen. However, there is a degree of overlap among these three curves for shorter-time activities. From Wells et al. (2009).\(^7\)

The differences between oxidative, phosphagen and glycolytic metabolisms are in how and how much ATP is produced, and how each pathway can be triggered by PA. In summary, the

\(^7\) Reprinted from Wells, G. D., Selvadurai, et al. (2009) Bioenergetic provision of energy for muscular activity. Paediatric Respiratory Reviews 10(3): 83-90. Figure 1\(^{st}\), with permission from Elsevier.
phosphagen system is normally triggered by activities that need immediate and intense muscle power (weight lifting, sprinting, or high jumping); deposits of ATP exist in the muscles but are depleted rapidly. Thanks to the phosphagen system, ATP is resynthesised from ADP and molecules of phosphocreatine. Running out of phosphocreatine forces the glycolytic and oxidative systems to take over in the production of ATP, with an inevitable drop in power output. Examples of activities that trigger the glycolytic system are repeated sprints in team sports or an 800-meter run; generally in these situations glucose from food is needed, and a complicated sequence of ten enzymatic reactions breaks it down to form molecules of pyruvate, with the production of ATP molecules (albeit in limited quantities). Pyruvate is further processed by the oxidative pathway. In longer-lasting physical activities (such as walking, running or cycling) oxygen is recruited from breathing and is involved in a complex set of chemical reactions divided into two big processes: the Krebs cycle and the electron transport chain. These two sets of metabolic steps lead to the production of a large number of ATP molecules and represent the most efficient pathway for energy production.

This description refers to what happens in muscular tissue but does not automatically suggest any systemic impact of any of these pathways. Of the three metabolic systems, the one that is reported to have greatest positive long-term impact on the whole human body is oxidation, and its benefit is identified in terms of enhancement of the cardiovascular system (Wells et al., 2009). It is therefore interesting to understand how aerobic activities affect the brain in the short- and long-term. Major focus on CE will be given in this chapter.

In the discussion section of their paper, Rovio and colleagues (2010) state that there are two hypotheses suggesting an effect of CE on the central nervous system: vascular and neurobiological mechanisms. The “vascular hypothesis” stems from the evidence of cardiovascular benefits of the type of exercise mentioned above, and speculates that the brain is positively affected by an aerobic enhancement because of the intricate network of vascular vessels that supply the cerebral tissue and because of the positive effect of exercise on risk factors for degenerative processes (e.g. cholesterol concentration). On the other hand the “neurobiological hypothesis” asserts that exercise would enhance a large set of mechanisms associated to neuroplasticity, such as neurogenesis and angiogenesis. These hypotheses are independent (Ahlskog, Geda, Graff-Radford, & Petersen, 2011), but there is supportive evidence from human and animal models for both of them. Most of the research carried out on human participants supports the vascular hypothesis as it is based on cardiac and blood-related measures. On the other hand, there is not a great deal of evidence supporting the neurobiological hypothesis; most of the evidence for this hypothesis comes from animal
models, as the methodologies are highly invasive\textsuperscript{8}. With regards to the neurobiological hypothesis, supportive data can be generally collected only post-mortem, as they require lab-assessment of parameters of the central nervous system. However, for the purpose of the upcoming review, a human-centred, rather than a hypothesis-centred approach will be used.

The theoretical scenario suggests that PA may have a beneficial effect on general brain health and may therefore play a protective role against detrimental factors attacking the brain. For this reason, engagement in PA has been investigated as a possible significant predictor for onset of dementia and neurocognitive decline. As oxidative metabolism has been reputed to be responsible for most benefits of PA, importance will be given to studies describing activities based mainly on oxidative metabolism. Although the number of published papers is large, no interpretational agreement has been reached, as such a comprehensive review is needed to include as much research as possible, to understand what exactly is the link between PA and neurocognitive variables. None of these papers can boast a theoretical model that goes beyond the cardiovascular/neurobiological hypothesis, and for this reason there have been no homogeneous methods of studies, which in turn leads to the expectation of an extremely variable pattern of findings.

2.1.2 PHYSICAL ACTIVITY, BRAIN AND COGNITION IN OBSERVATIONAL DESIGNS

2.1.2.1 USE OF QUESTIONNAIRES

A considerable number of studies have investigated PA and the likelihood of developing dementia in large cohorts of individuals, who were followed up for a certain amount of years. In this kind of research PA has been used as a predictor to account for conversion from normal cognition to dementia. The main strengths of this kind of research are the number of participants (usually thousands of individuals) and the longitudinal design, which can monitor possible decline over time. Unfortunately there are also weaknesses: first, the engagement in physically stimulating activities has been often measured through questionnaires: these are reported to be valid and reliable, but yet are only an indirect measure of the true amount of

\textsuperscript{8} No biological model was detailed in this review, if carried out in other species than humans. This has been done on purpose: ethically-wise, animal research gives more chances for experimental manipulation in vivo, and allows the researchers to control for a huge number of variables, which is impossible with humans. On the other hand, it can be argued that humans display a greater degree of complexity than any other animal model. For this reason any findings obtained for example in a mouse model should be extended to humans with great caution (Roses, 1997). Since the mechanisms of action associated with AD pathology are today largely unknown, it would be hazardous to extrapolate results obtained in a mouse model to humans.
physical exercise done by the individual; second, there has been no agreement on the activities that can be considered potentially beneficial. There are activities that have been always included in the questionnaires, like walking or cycling, but then each team of authors has also included their own set of activities in the list, unfortunately, without making an informed choice based on a specific metabolism-based rationale. For this reason a team may have mentioned an activity that is not considered by others, and vice versa. Leisure activities have been sometimes included, such as dancing (Scarmeas et al., 2009), gardening (van Gelder et al., 2004) and carpentry (Abbott et al., 2004). Third, each team referred to their own idiosyncratic set of criteria with regards to intensity and frequency; fourth, there is variability in test-retest interval duration and in the baseline age of the sample; fifth, Coley et al. (2008) added other two factors associated with the link between PA engagement and its effects: the window of exposure and duration of benefit. No studies have formally tackled these two aspects, but, although being quite uniform, it might still be possible that they play an important role.

Large cohorts of elderly individuals were recruited when their cognition was unimpaired, and a questionnaire was administered to them to obtain a rough measure of the amount of PA they engaged in. They were then assessed again at retest after decades to determine the proportion of individuals who had converted to AD. Participants were usually split in two or more groups according to the amount of PA done (sometimes a metabolic conversion was computed), and group differences were tested. Some of these articles described the use of the standard criteria for diagnosis of probable AD (Mckhann et al., 1984). In these studies higher levels of PA are generally associated with decreased odds of developing AD: 5-6 (Larson et al., 2006; Laurin, Verreault, Lindsay, MacPherson, & Rockwood, 2001; Scarmeas et al., 2009) and even 20-30 (Andel et al., 2008; Rovio et al., 2005) years later. This association was not found in other studies (Abbott et al., 2004; Verghese et al., 2003).

Apart from one case (Scarmeas et al., 2009), PA was also reported to be a significant negative predictor for a general condition of dementia. In a particular instance, PA was reported to be protective against dementia, but it was beneficial to AD only in the sub-cohort of participants that had poor physical function at baseline (Taaffe et al., 2008). In another study, no association with AD or dementia was found (Ravaglia et al., 2008).

All reported results in these studies were “filtered” through the inclusion of a series of variables inputted as covariates: these included demographics, health variables, factors related to general physical function, lifestyle variables, and, sometimes, presence/absence of the ApoE $\varepsilon_4$ allele. Unfortunately there is a degree of variability in the choice of these covariates, and for
this reason, together with the aforementioned bullet points, the comparability of these studies is low. Deaths and drop-outs may have also played an important role. Generally speaking from a qualitative point of view, the conclusion emerging from these studies is that PA may have a beneficial impact against AD and dementia, but an observational study is probably not the best methodology to capture it.

Other studies with comparable methodology reported undefined levels efficacy of PA on cognitive maintenance against the onset of MCI: PA has been found to be protective against a diagnosis of “Cognitive Impairment, No Dementia”, but not against a diagnosis of “amnestic non-vascular MCI” (Middleton, Kirkland, & Rockwood, 2008). In a parallel study only light or vigorous exercise intensity was found to be effective in comparison to moderate activity, and only when it was performed once or twice a week, or 5-6 times a week, but not when it was daily or performed three or four times a week (Geda et al., 2010). It is evident that these results do not depict a clear trend and are not helpful to draw clear conclusions.

Similar studies measured cognitive decline with no specific progression to dementia/MCI. In studies qualitatively similar to those described in previous paragraphs, longitudinal measurements revealed worse cognitive decline in the quantiles of healthy adults who exercised least (Etgen et al., 2010; Middleton, Mitnitski, Fallah, Kirkland, & Rockwood, 2008) or displayed a recent drop in the duration of PA they were used to engage in during old adulthood (van Gelder et al., 2004). General decline in cognitive tests and screening instruments was steeper also for adults with milder regimes of activities estimated at 2 years (Weuve et al., 2004) or 10 years (Jedrziewski, Ewbank, Wang, & Trojanowski, 2010) prior to assessment, or even at middle age (Singh-Manoux, Hillsdon, Brunner, & Marmot, 2005), or during early life (Dik, Deeg, Visser, & Jonker, 2003). In one case no association was found (Sturman et al., 2005).

These observational studies suffer from all aforementioned shortcomings but display more convincing evidence that PA triggers some sort of positive impact on cognitive abilities measured at a later age, although there are sporadic negative findings that should not be neglected (Coley et al., 2008).

A set of studies investigated cognitive functioning of smaller groups of unimpaired or undemented old adults as a function of PA in a cross-sectional fashion. Group membership to the higher quantiles of self-reported PA was associated with higher scores of inductive reasoning (Perrot, Gagnon, & Bertsch, 2009), executive functions and aspects of memory (Eggermont, Milberg, Lipsitz, Scherder, & Leveille, 2009), short- and long-term memory, general cognition and semantic skills (Lam et al., 2009). Only a trend for executive functions
was found by O'Dwyer and colleagues (2007). Generally speaking, these studies do not add relevant evidence to the overall picture, and it is really challenging to draw a set of homogeneous conclusion with the inclusion of all the peculiar and unique aspects of each individual study.

2.1.2.2 USE OF PHYSIOLOGICAL-MOTOR MEASURES

There is research that has aimed to interpret the relation between cognition and PA using more objective variables: indices of cardiovascular or cardiorespiratory fitness. These indices are values measured (or estimated) with specific accurate tests monitoring physiological variables like venous/arterial blood oxygen concentration, Heart Rate (HR), and gas concentration during inhalation/exhalation associated with exercise. There is proof that PA improves fitness (Baker et al., 2010), but it is likely that fitness cannot be accounted for solely by PA. For this reason, studying the relation between fitness and cognition would be informative but not very relevant for the purposes of this review. Nonetheless, since PA is linked to the vascular hypothesis, and since the degree of correspondence between the two measures was reviewed as strong (Oja, 2001), it is advantageous to broaden the relation between cardiovascular fitness and cognitive skills. The most used parameter of fitness has been the rate of maximal oxygen consumption (VO$_2$max), a physiological parameter usually obtained through a graded exercise test, which gives an estimate of the general condition of the individual on their capacity to engage in aerobic activity.

Attention, executive functioning and global cognition (Netz, Dwolatzky, Zinker, Argov, & Agmon, 2011), and spatial memory (Szabo et al., 2011; Voss, Erickson, et al., 2010) were positively correlated with VO$_2$max but not with the self-reported score of PA, suggesting that PA itself might be unrelated to cognitive health, whereas fitness levels is. Fitness levels of younger adults were also associated with the whole cognitive sphere 6 years later (Barnes, Yaffe, Satariano, & Tager, 2003). This pattern is not as simple as it may seem, since performance in a verbal learning test was not reported to be associated to fitness but was associated instead to the metabolic conversions of PA reports (Floel et al., 2010). Higher attentional, working-memory, and speed-of-processing levels were found in highly-fit adults compared to low-fitness adults, and an age-by-fitness interaction revealed that the difference was strangely steeper for the young subgroup (Newson & Kemps, 2008). This suggests that the impact of age in the PA-cognition relation might be partially spurious, and might mask a pure effect of baseline fitness. As supportive evidence of this view, VO$_2$max was linked to fewer perseverative errors on the Wisconsin Card-Sorting Test (WCST), regardless of age (Voss, Erickson, et al., 2010).
It is possible that the link between PA and cognition is not exclusively explicable as a function of cardiorespiratory fitness: recent studies found positive associations between cognitive abilities and VO$_2$\textsubscript{max}, grip force, and motor fitness alike, indicating that parameters like fine coordination, balance and flexibility play a role (Voelcker-Rehage, Godde, & Staudinger, 2010). The presence of dementia was reported to be also linearly related with simple rates of mobility (Prince et al., 2011) and AD onset was also associated with general physical function (Wang, Larson, Bowen, & van Belle, 2006), muscle strength (Boyle, Buchman, Wilson, Leurgans, & Bennett, 2009), and walking speed (Marquis et al., 2002). Gait speed also explained a significant portion of the variance related to Stroop and TMT-B performance in a sample of MCI (McGough et al., 2011). These aspects seem to be clinically important because they can be easily collected for simple screening purposes years before a diagnosis: however for the purposes of the review they will be left out.

2.1.2.3 OBSERVATIONAL STUDIES AND NEUROIMAGING

2.1.2.3.1 BRAIN STRUCTURE

Many techniques have been used to study the link between brain structure and engagement in PA, and various diagnostic groups have been investigated.

A positive association was found between fitness estimated with a 6-minute walk and hippocampal and entorhinal volume in a sample of amnestic MCI. On the contrary, measures of physical function were uncorrelated (Makizako et al., 2011). A similar positive association was detected in a sample of CDR=0.5-1 adults with a diagnosis of probable prodromal AD: VO$_2$\textsubscript{max}-fitness was positively associated with total GM and WM volume (Burns et al., 2008), and with the total size of every single lobe, with clusters surviving Family-Wise correction located in the inferior parietal cortex bilaterally (Honea, Thomas, et al., 2009). A specific focus on the volume of the small mediotemporal structures revealed that this association was valid for the parahippocampal complex (Honea, Thomas, et al., 2009).

A larger number of studies investigated the correlation in healthy individuals. No associations were found with total GM and WM (Burns et al., 2008) or with WM hyperintensities (Podewils et al., 2007). The mediotemporal formation was reported to be uncorrelated to fitness (Honea, Thomas, et al., 2009), whereas the hippocampus has been reported to be correlated (Erickson et al., 2009). Another study replicated this last finding but found no association between hippocampal size and scores of self-reported PA (Szabo et al., 2011).
Additional explorative analyses have been published with variable findings. Registration of VO2max and voxel-based MRI scans at the same timepoint was carried out by Gordon and his team (2008) in a mixed sample of young and old adults. They reported an association between fitness and GM in mediotemporal, anterior parietal and inferior frontal clusters, but no differences were evident for WM. In a sample of non-demented elderly individuals TBM revealed that the areas involved were the occipital-fusiform areas and the lentiform nucleus only (Honea, Thomas, et al., 2009).

Despite the lack of association between PA measured via questionnaires and hippocampal size (Szabo et al., 2011), weekly estimates of PA-dependent metabolic consumptions correlated with regional brain volume in the prefrontal cortex, cingulate gyrus, occipitotemporal clusters and the cerebellum (Floel et al., 2010). In another study old adults with higher levels of PA had larger temporal WM, especially around the parietooccipital junction (Ho et al., 2011).

There are some teams that chose instead to consider PA as a measure of activity done earlier in life. Self-reports of a cohort already described in other studies (Kivipelto et al., 2008; Rovio et al., 2005) revealed an association between higher engagement in PA at middle-age with larger GM volume 20 years later, with no influence on WM (Rovio et al., 2010). Moreover, those who had exercised at least twice a week displayed increased GM in the middle frontal gyrus. In another study, middle-age-to-old healthy participant with higher levels of activity over the latest 10 years had larger volumes in superior frontal and pericalcarine cortices. In addition, an age-by-PA interaction was solely found in a mediotemporal ROI comprising hippocampus, parahippocampal gyrus and amygdala, where significantly larger atrophy was found for inactive old adults (Bugg & Head, 2011).

A few studies investigated parameters of structural connectivity: ROI-DTI was performed in a mixed sample of young and old adults and the correlation between metabolic conversion of reported PA and FA was only significant in the uncinate fasciculus and in the cingulate region, with no frontal involvement (Marks et al., 2007). A recent study looked for the link between WM integrity and PA in an exploratory design and fitness levels accounted for increased FA and decreased radial diffusivity in healthy middle-age-to-old adults in a portion of the corpus callosum (Johnson, Kim, Clasey, Bailey, & Gold, 2012).

2.1.2.3.2 BRAIN FUNCTION

Evidence of the effects of PA or fitness on brain function at rest is scarcer, and most of the times it has reflected paradigms of neuroimaging with a task (Colcombe et al., 2004; Rosano et al., 2010). To my knowledge, scarce interest has also been shown in other pathologies as well
The only relevant study based on a resting-state paradigm was published recently (Voss, Prakash, et al., 2010). VO\textsubscript{2}max of young and old adults was registered and the PCC was selected as seed region; the areas with age-dependent differences in connectivity were chosen, and the connections between these ROIs were tested together with the fitness parameter in a correlational analysis. The only correlation surviving to Bonferroni correction was that between VO\textsubscript{2}max and the connection from the seed ROI and the middle frontal gyrus. To date there is no study focusing on the relation between PA and brain function at rest in AD.

### 2.1.2.3.3 PEPTIDIC MARKERS

As a final remark, the link between measures of PA/fitness and AD molecular biomarkers has been studied only in few occasions in humans. Liang et al (2010) measured PiB deposition rates, CSF biomarker concentration and PA in a sample of CDR=0 older adults and split the group according to various cut-off scores documented in the literature: subgroups displaying higher CSF p/t-TAU concentration/PiB staining and lower CSF A\textsubscript{β\textsubscript{42}} concentration at the same time also displayed a significantly lower amount of self-reported PA. The same pattern of results was found when inverting dependent and independent variables. However, none of these findings was still significant after adjusting for demographics. Recently, the same authors tried to investigate the same variables with an inverse methodology: PiB staining and CSF A\textsubscript{β\textsubscript{42}} were analysed as a function of PA: main effects of the predictor were found on both dependent variables, with an additional interaction effect only expressed on in vivo Amyloid deposition (Head et al., 2012). Some authors have suggested that biomarkers' level could be used as indices to evaluate the efficacy of exercise (Foster, Rosenblatt, & Kuljis, 2011), but no attempt has been formally made yet.

The picture of findings is quite confusing, as it is impossible to extrapolate a common trend. One possible explanation is in the nature of the measurement of PA: a variable that implicitly contains several methodological aspects of variability is destined to give a variable pattern of results, which prevents from any chance of interpretation. Trying to draw an overall conclusion from observational studies, there is some evidence that engaging in chronic PA attenuates neural and behavioural parameters of cognitive decline. Looking at the overall picture, the effects of PA do not appear to affect brain and cognition in a global and uniform fashion; there is instead a trend suggesting that the main influence is probably not exerted on the temporal lobe, but rather on the anterior half of the brain and to functions mainly based on executive control.
Part of the picture links the benefit of PA to an enhanced level of cardiorespiratory fitness, which supports the cardiovascular hypothesis. Research testing the neurobiological hypothesis has mainly focused on animal models, because animal research allows the experimenters to measure variables that are of hard access in humans (and are often assessed post-mortem). Moreover, it has intrinsically been associated with procedures in which variables of interest are manipulated rather than observed.

Even though observational research leads to somewhat undetermined results, it is still possible and of scientific interest to study the impact of PA, but it is necessary to move to branches in which a direct manipulation of this variable has been performed.

2.1.3 PHYSICAL ACTIVITY, BRAIN AND COGNITION IN EXERCISE INTERVENTIONS

2.1.3.1 INTERVENTION STUDIES AND BEHAVIOURAL VARIABLES
Trials testing the efficiency of a protocol of PA display important strengths compared to observational studies. In the first instance it is not as time-demanding as following a large cohort over decades, since programmes only last weeks or months; second, the experimenter is fully aware of the variables that have most likely played a role in the training. There is a more precise measurement of PA in terms of activities done, regime intensity, weekly frequency and session duration, and confounding factors are less likely to influence the results; third, test-retest measurements are possible, and the resulting difference-scores are cleansed of further intervenient variables. Unfortunately, weaknesses are also visible: rarely have two paradigms used similar parameters. This partially compromises the direct comparison of the results found by two teams. For this reason it is necessary to generate an interpretational frame in which to insert all findings with caution. Participants in PA programmes must normally have a high degree of commitment and be willing to complete multiple assessments and a training regime that might turn into a heavy burden. For this reason, samples are not always large, and rates of dropout must not be neglected.

2.1.3.1.1 INTERVENTIONS IN DEMENTIA
Compared to observational research, there are a limited number of papers reporting the effects of a PA programme administered to a sample of demented patients. Usually the interest is in basic outcome variables, such as ADL, quality of life, undesired psychiatric or behavioural symptoms, general health enhancement, variables linked to the physical domain, and benefits for caregivers. The rationale supporting the feasibility of these programmes in an
AD sample comes from the evidence that motor areas remain relatively devoid of AD pathological markers even at latest stages of the disease (Braak & Braak, 1991). Exercises based on simple movements (like walking for instance) can be easily implemented in this population, and benefits may be expected. Two meta-analyses searched for any trials that aimed to improve performance on behavioural measures in AD through protocols of PA. Heyn and colleagues (2004) included 30 studies in their analysis, concluding that physical interventions objectively improved cognition in dementia. This apparently convincing conclusion is however obscured by methodological ambiguity. Outcome measures focused mostly on cardiovascular, strength, functional and behavioural variables and only 8 studies out of 30 actually included a cognitive outcome (all based on an intervention aimed to trigger an oxidative metabolic response). Moreover, the initial MMSE score of the participants was not available in one of those papers, and it varied from 6 to 25 in the remaining 7, indicating no uniform pattern of severity for the tested samples. Modality of exercise was also highly variable, with a variable number of weekly sessions and a programme duration ranging from 2 to 16 weeks. A more recent meta-analysis used stricter inclusion and exclusion criteria (Forbes et al., 2008). The authors stringently reviewed any candidate study and concluded that only four papers reached acceptable standards, three of which did not include any cognitive outcome. As expected, no benefit from PA was evident from this meta-analysis. These two papers studied the effects of PA over a general diagnosis of dementia, without specific highlight on AD. Yu (2011) included in her review studies whose sample included AD patients only. Despite the high inter-research variability in study characteristics, she concluded that there is optimistic evidence of an impact of PA, at least on general measures of cognition like MMSE. Not all interventions she reviewed were exclusively based on PA, though, as she incorporated two papers with mixed interventions (exercise + cognitive activities).

Focusing on each single relevant paper might help draw a picture of the effects of PA trials on AD diagnosis. Three short tests were administered to a sample of 27 moderate patients before and after either 2 sessions of home safety training or a 6/12-week home-based mixed protocol of brisk walking, and training of strength, flexibility and balance. None of the cognitive scores exhibited improvement (Steinberg, Leoutsakos, Podewils, & Lyketsos, 2009). Walking was also the core of the attempt made by Eggermont and colleagues (2009), who recruited patients with an MMSE score between 10 and 24 and invited them to walk for half an hour for 6 weeks, 5 times a week. No test-retest differences were found between patients and controls (who were just receiving social visits) in the cognitive domains. These two studies suggest that walking is perhaps not sufficient to trigger a positive benefit in cognition when the pathology is full-blown. However, in a third study, a 15-week programme of walking, balance training and
stamina exercises elicited a significant improvement in a French MMSE-like screening tool, with a concurrent decrease exhibited by the controls (Kemoun et al., 2010). The two groups were matched by MMSE score (even if highly variable), but unfortunately, the authors did not describe what the activities of the control group consisted of, nor was the classic NINCDS-ADRDA criteria used to diagnose AD. Tougher protocols have rarely been attempted: Palleschi and colleagues (1996) ideated a cycling machine programme, 3 days a week for 3 months, in which their AD participants (MMSE: 18-21) were instructed to cycle at 70% of their HR for 20 minutes each time. Significant improvements were registered in MMSE, digit cancellation, verbal span and supraspan. This study, albeit old and without a control group, is interesting because the authors performed a pharmacological wash-out before starting the treatment. An AD drug therapy may be an important confounding factor in this kind of designs for both methodological reasons regarding the sample, and the nature of the experimental manipulation. In the first instance policies regarding medication vary from country to country, and this contributes to making the comparison between studies difficult; second, variability in medication type and dosage may exist within the same sample; third, response to drug treatment may differ from person to person; fourth, what is believed to be the positive effect of PA may be instead the synergic effect of PA plus drug treatment. Unfortunately, apart from the mentioned study, specific descriptions of medication intake were not included. In conclusion, there is a need for further evidence from protocols of tougher intensity, and in this respect, future research has already been planned (Cyarto et al., 2010; Pitkala et al., 2010). As final mention, a team (Van de Winckel, Feys, De Weerdt, & Dom, 2004) assigned AD patients either to face-to-face conversation sessions with a therapist or to dance-like mimicking sessions supported by music. After 12 weeks, MMSE and category fluency scores improved in the experimental group whereas no change was reported for the controls. The authors referred to this treatment as a combination of music and exercise, but there appears to be no sufficient grounds to equate this type of intervention to protocols based on walking or heavier PA. Unless one recognised the strength of the treatment in the mimicking component (thus classifiable as cognitive stimulation, rather) or in the music itself, it seems unrealistic that such a programme might be sufficient to trigger a cognitive improvement. This prototypical study with a small sample highlights the fact that better studies are needed.

2.1.3.1.2 INTERVENTIONS IN MCI

Only a few trials have aimed to improve neurocognitive measures in samples of MCI through PA: similarly to dementia, a Cochrane review has been planned and will be published in the future (Orgeta, Regan, & Orrell, 2010). Also, experimental trials have been designed and the data collection is currently on-going (Liu-Ambrose et al., 2010). Five main studies have been
identified. One-year walking at moderate pace was the core form of the treatment for a Dutch sample of MCI (van Uffelen, Chinapaw, van Mechelen, & Hopman-Rock, 2008). Cognitive abilities were compared to a control condition in which activities such as posture, balance, and relaxation exercises were performed. No between-groups differences were found, confirming the idea that walking is not enough to trigger a neural response as a form of CE. However, this evidence is not supported by the findings of an earlier study: Scherder et al. (2005) recruited a sample of old frail MCI patients to test the efficiency of a 6-week-long programme of self-paced slow aid-assisted walking (30 minutes for 3 times a week), compared to a control group receiving either social or no contact. In the experimental group significant improvements were observed in the TMT-B and category fluency tasks, with no changes in memory. Attempts have been made also with heavier training regimes. A multivariate ANOVA was carried out for executive and memory functions before and after six months of either heavy aerobic workout or stretching exercises: MCI participants engaged in their set of activities 4 times every week for about one hour; the treatment resulted in executive improvements but no changes in memory (Baker et al., 2010). Lautenschlager and colleagues (2008) recruited a large cohort of participants with subjective memory complaints and objective impairment (without using the official MCI criteria, though), allocating them either to a control condition based on educational material or an experimental condition in which they were asked to exercise three times a week for 50 minutes, engaging in moderate activities. Results showed that exercisers improved in the ADAS-COG battery and in verbal long-term memory, with no impact reported on short-term recall, verbal fluency and DSMT.

Overall, these studies do not draw a uniform picture, and there is only weak evidence that when activity is not heavy enough it does not cause any benefit. Recent evidence showed that two sensibly different training regimes brought equal benefit to MMSE scores in MCI patients (Varela, Ayan, Cancela, & Martin, 2012), but when specific tests are used, more challenging programmes appear to be beneficial. However, it is yet not clear which cognitive component is the core of the improvement.

2.1.3.1.3 INTERVENTIONS IN HEALTHY INDIVIDUALS

The first meta-analysis of studies carried out with healthy adults was performed by Colcombe and Kramer (2003). They included fourteen longitudinal trials published between 1966 and 2001 and reported a benefit of exercise on all investigated domains, with the highest significant effect found for improvements in executive functions. A Cochrane review (Angevaren, Aufdemkampe, Verhaar, Aleman, & Vanhees, 2008) included a meta-analysis of old studies (the most recent one was published in 2002) based on questionnaires and fitness
levels, selecting the papers with strict inclusion and exclusion criteria. Out of eleven studies, eight reported an improvement in fitness and nine reported a positive impact on at least one cognitive aspect. However, the cognitive domain influenced by PA was not constant and the majority of the $p$ values were not significant, although measures of processing speed and attention seemed to be the aspects benefitting most. These two articles demonstrate once again how inclusion and exclusion criteria play an important role when it comes to meta-analysis in an area of research characterised by extreme variability. Etnier et al. (2006) hypothesised that a relation between fitness enhancement and cognitive improvement would be displayed by a dose-response effect. Thirty-seven papers dating up until 2004 were considered and different sections of the paper targeted cross-sectional and pre-post comparisons. No linear relationship was found in the first branch of studies whereas a significant negative regression line was fitted in longitudinal designs. Once again, however, the variability of the included studies was very high. Again, a glance on more recent studies can be helpful. “Nordic walking” training (consisting of normal walking aided with poles to train the upper body) was compared to a no-contact and a stretching/flexibility control groups: after 6 months (with a regime of at least 3 weekly sessions) a 3-sample ANOVA analysing memory levels was non-significant, but when the two exercising groups were taken together, an improvement was registered (Ruscheweyh et al., 2011). However, the findings of this study are enigmatic, as the flexibility group (strangely defined by the authors as “low-intensity aerobic exercise”) obtained better memory scores at retest. A 1-year combination of aerobic training and other activities was piloted by Williamson et al. (2009) without the inclusion of a control group. A significant improvement was found in the DSMT after treatment, but the performance in the Stroop test and in the Rey Figure did not change. A dose-response was sought by Masley, Roetzheim & Gualtieri (2009), who tested a sample of extremely variable age before and after ten weeks of fitness-improving exercise done either 3-4 or 5-7 days a week. Only cognitive flexibility benefitted and a dose-response was also evident when comparing the two training regimes. The authors concluded that executive skills were the only functions sensitive to exercise. The effect of twelve weeks of stretching or cardiovascular exercises was explored in relation to the WCST by another team (Albinet, Boucard, Bouquet, & Audiffren, 2010). A group-by-timepoint interaction was found; three weekly sessions of aerobic activities translated into an average decrease of ten errors for the exercisers and led to a 2-error increase for the controls. A Spanish study recruited older women and trained them intensively (5 times a week) for five months with water-based (reasonably aerobic) training combined with either calisthenics or strength exercises. In both groups an improvement was recorded in the MMSE, but unfortunately, no pure control group was used (Carral & Perez, 2007). In another study, ten months of cardiovascular workout of increasing intensity resulted
in improved performance in the Stroop Task only, with no effects on the WCST and measures of pure RT. The control group did not show the same benefits, and these selective improvements were not associated with enhancements of VO$_{2}$max level (Smiley-Oyen, Lowry, Francois, Kohut, & Ekkekakis, 2008). An on-going ambitious project is being carried out in Finland by Komulainen and colleagues (2010) who have been following up more than twelve hundred healthy elderly divided into six different intervention groups. Two of these groups are relevant for our review as they focus on aerobic exercise. The first six months of intervention were shared by both groups, and consisted of a gradual increase in load intensity; after that, two subgroups were created according to the weekly amount of energy expenditure of subsequent training period. As the participants were simply instructed to follow a training regime but were not explicitly guided during the training, a questionnaire was administered to them, and VO$_{2}$max was also measured. Results showed group membership led to no change in cognitive functions, but change in VO$_{2}$max positively predicted performance in immediate and delayed memory scores of the CERAD battery. A Japanese study (Nakayama et al., 2011) tested older adults with the 3MS, the MMSE, and the Frontal Assessment Battery (FAB) before and after 12 weeks of daily moderate bench stepping (twice a day, 140 minutes a week in total). The participants were divided into independent and semi-independent adults and assigned to the experimental or control group. Training improved scores in all three batteries in the semi-independent subgroup, whereas independent adults obtained benefit in the 3MS and MMSE only (possibly because of a ceiling effect for the FAB, as inferable from the graph). However, these conclusions were drawn after performing 4 separate t-tests (one per subgroup), rather than a comprehensive analysis, and for this reason the results should be interpreted carefully. A different attempt was recently made in Germany, where middle-aged adults were assigned to no, stretching/coordination, or cycling training in a very good design: the aerobic group trained for six months at a hard intensity and changes in VO$_{2}$max were registered together with changes in various cognitive functions; both interventions resulted in improvements in learning and the stretching/coordination group overcame the aerobic group in attention scores, with no effect detected in executive functions (Hotting et al., 2012).

The teams responsible for all these trials have preferred the recruitment of older participants for various theoretical and methodological reasons, and have not opted for a sample of young adults. Griffin and colleagues (2011) specifically enrolled young adults to test the effects of 3 or 5 weeks of moderate exercise performed for 60 minutes three times every week. They were tested using a face-name association task before and after the programme, and improvements were recorded for the exercisers only. The interesting aspect is that these individuals, despite being young, were reported to be untrained: young adults are easily fitter than older adults,
but it appears possible that baseline fitness becomes a key-aspect when age is homogeneous in the sample. This consideration will turn to be important in Chapter 4.

The body of research studying the effects of specific PA programmes suffers from the shortcomings previously discussed; however, a general trend suggests that measures of executive functions seem to respond positively to either increase in fitness or simple enrolment in the experimental group.

2.1.3.2 INTERVENTION STUDIES AND NEUROIMAGING

2.1.3.2.1 BRAIN STRUCTURE
Structural changes after an exercise intervention were addressed by the team of Colcombe (2006), who reported the effects of a 6-month protocol of activities aimed to improve cardiovascular fitness in a sample of older adults; fitness was augmented in the experimental group as expected, and an increase in GM volume was observed in ACC, inferior and middle frontal gyri and the supplementary motor area bilaterally, with an extension to the left superior temporal lobe. A WM change was also found significant in a cluster located anterior to the genu of the corpus callosum. Moreover, no volume changes were documented in a young (18-30 y.o.) control group, and this interestingly clashes with the findings by Gondoh and colleagues (2009), who specifically focused on a sample of young adults. In this study 4 months of aerobic dance classes (at least one per week) led to maintenance of GM volume, while in the no-activity control group, a decrease in left insular volume was reported. As the hippocampus is too small for exploratory analyses, a ROI was specifically drawn by Erickson and his team (2011). They recruited a very large sample and assigned their participants to an aerobic or stretching-toning programme. Group-by-timepoint interactions were found for hippocampal size and cardiovascular fitness (both increased in the experimental group).

Pereira et al. (2007) used blood volume changes as parameters of estimated adult hippocampal neurogenesis. A small sample of young participants completed a 3-month programme of aerobic training, and their hippocampal maps were investigated at baseline and retest. The dentate gyrus was the only structure displaying an increase in blood volume, correlated with VO\(_2\)max. As the dentate gyrus is the only subfield where neurogenesis still happens in adult life, exercise triggered a plastic response in this structure particularly sensitive to ageing and pathological processes.
Again, no certainty can be extracted from studies on brain structure, but the general conclusion suggests that exercise can affect variables of brain structure, albeit mechanisms and target areas are not clear.

2.1.3.2.2 BRAIN FUNCTION

In the second part of their paper, Colcombe et al. (2004) described the effect of a 6-month intervention, in which older adults trained for three times every week following a programme of aerobic exercises of increasing challenge. Compared to the stretching/toning controls, the experimental group obtained better behavioural scores in an Eriksen Flanker test and showed enhanced activation during this task in frontal-parietal areas, with a decrease in the ACC. In a completely different design, a 12-month intervention was administered to older adults, with focus on walking exercises of increasing load; functional connectivity was measured at baseline, seeded at PCC and limited to pre-determined ROIs on which ageing processes were known to have an effect. No change in connectivity was recorded when the training and the control group were compared with a MANOVA, but differences emerged after running separate univariate ANOVAs. The training group displayed increased connectivity in the mediotemporal lobe, between temporal and prefrontal regions, and between temporal and occipital areas. This last component was correlated with an improvement in executive functions (Voss, Prakash, et al., 2010). In summary, these two disparate studies suggest that exercise interventions can positively alter brain function in some way.

2.1.4 CONCLUSION

An extensive literature has studied the effects of physical activity on brain and cognition. Most of the evidence comes from observational designs and only a limited amount of randomised trials have been ideated and successfully completed. In comparison, far too many reviews, editorials, letters and “calls to arms” have been published, most of them report quite too optimistic conclusions, often limiting the supporting evidence to just a few studies and ignoring those findings that are not pertinent with a positive trend. There is theoretical potential for a powerful effect of exercise, but unfortunately there is too little evidence and too many intervenient variables. Methodological factors may account for the big variability found in the results, especially in the studies that have investigated neural variables. The main problem with this topic is that at the moment the cardiovascular and neurobiological hypotheses are not satisfactory. They are good at predicting the global physiological changes that may be triggered by PA, but they do not suggest anything regarding what specifically
happens in specific regions of the brain, particularly in relation to what it is possible to detect measuring brain function and structure. A general overview of the papers reporting a positive impact of PA would suggest that the benefit is prominent in frontal areas and executive skills, and this could be considered a good starting point for the formation of specific working hypotheses. As nothing is known about neural processes associated with PA, exploratory neuroimaging studies are preferred to the use of ROIs, and more studies investigating parameters of connectivity are needed. An overall reasonable consideration is that investigating cognitive measures does not have priority. PA is a mechanism that does not require any cognitive computation but causes a series of “physical” or “organic” effects on the body. It is therefore likely that any benefit of PA on cognition will be spurious, and rigorously modulated by the nervous system. For this reason more focus should be reserved on analyses of brain function and structure, and only secondarily on cognitive functions. At the moment no theoretical model with neural predictions exists for CE. Vice versa, a simple and theoretically appropriate model will be described in Chapter 4 to account for the effects of AE on brain function, and it will be possible to observe how cognitive evidence assume a completely different meaning in the light of a strong pattern of support. In addition, it is worth considering that a “physiological” variability in exercise-response is naturally present in a population of untrained people (Hotting et al., 2012), and this may sensibly contribute to the lack of homogeneity in the findings.
2.2 COGNITIVE ACTIVITY

2.2.1 INTRODUCTION

Engaging in cognitive-stimulating activities has been reputed protective against the onset of age-associated cognitive decline and abnormal ageing. Similar to the framework described for PA, CA could exert effects on neural substrate and cognitive abilities. PA and CA can be operationalised in a similar way: both can be manipulated in terms of difficulty, duration and frequency, and both can also be conceptualised as “baseline fitness”. The idea of physical fitness is quite straightforward, but the notion of “mental fitness” (introduced by Salthouse (2006)) is not. Jaeggi et al. (2011) suggested that working memory is the cognitive correspondent of physical fitness, since it is necessary to perform a big variety of tasks. However, a reasonably better incarnation of cognitive fitness exists: it could be considered as part of the available “cognitive reserve” (CR), in turn strictly associated to the concepts of “brain reserve” and “neural plasticity”.

Brain reserve refers to the capacity of the brain to recover functionally from neural insults (Stern, 2009). Any process of recovery depends on the amount of healthy neural substrate available for processing, and thus on quantitative measures such as neuronal number and brain volume. This mechanism is passive, because the nervous system accomplishes it automatically. CR refers instead to the active process the brain engages in to cope with neural insults by using compensatory or alternative cognitive elaborations (Stern, 2009). Brain reserve and CR take part in neuroplasticity as they protect against brain damage and degeneration.

Variables such as baseline IQ level, years of education, occupation, and social extraction influence the level of CR (Valenzuela & Sachdev, 2006), and unfortunately CR cannot be explicitly measured, but only estimated. Cognitive, behavioural and even neural variables have been used as indicators (Steffener & Stern, 2012), and the hypothesis that a neural variable can be used as an estimate of CR (rather than of brain reserve) comes from the idea that CR needs some form of neural implementation. Stern (2009) has investigated this neural component and has distinguished between the ideas of “neural reserve” and “neural compensation”. Neural reserve refers to the degree of individual difference in brain function of healthy adults during task performance. This variability can reflect different levels of neural efficiency the involved networks are capable of, and this efficiency is supposed to play a role in task management, as task difficulty increases. This aspect will be important in one of the upcoming chapters as it may play an important role in paradigms of task-associated fMRI recordings. Neural compensation is involved instead in coping with brain pathology. If a task-
related network was damaged, an alternative computational circuit would be recruited, often without allowing the patient to reach a standard-level behavioural performance. According to this model, the process of healthy ageing would tap brain efficiency and neural reserve, whereas AD would trigger mechanisms of neural compensation. Both these processes are tightly linked with mechanisms of plasticity.

The idea of plasticity was introduced in the first chapter as the backbone of Mesulam’s (1999, 2000) theory of AD and mentioned in relation to the effects of CE, but it needs further clarification: Lövdén and colleagues (2010) proposed a model defining it as the overall ability of the brain to change reactively to a unbalance between resources available to the person and resources requested by an external input (e.g.: a cognitive task); the authors further discriminated between plasticity and flexibility, where flexibility indicates a change that is directly possible using the available neural supply, whereas plasticity assumes some underlying structural modifications inside the brain. Based on this, CA should trigger plastic mechanisms that can be measured through neural and behavioural examinations. In this light, cognitive estimates of plasticity (Calero & Navarro, 2007; Uttner et al., 2011) appear as weak theoretical constructs.

One of the main applications of plasticity models is the possible effect of CA on cognitive variables: when a person is intensively trained with a set of tasks, they may display gradual improvement in performance: with the training material; with the same tasks but based on different material; with untrained tasks tapping the same cognitive components as the training material; with different tasks tapping other cognitive aspects than the training material. These four categories are relevant to interpret the efficiency of the training, and are strictly linked to the underlying neural restructuring. “Practice effect” defines the first two categories, whereas “transfer” identifies the other two. A further subdivision distinguishes between “near” and “far” transfer. When improvement is recorded in a different task, it means that a transfer of cognitive gain took place. In most cases transfer effects are reported in tasks that are similar to or share computational elements with the training material (near transfer); sometimes the benefit is extended to a task that measures a distant cognitive function (far transfer). Near transfer is commonly observed but, on the other hand, far transfer is rare. Barnett and Ceci (2002) identified numerous factors that take part in a comprehensive model of transfer mechanisms; their model, however, is explorative and completely based on behavioural evidence. An alternative view considering neural aspects as well might be more useful in understanding how transfer works. In all likelihood improvements triggered by training are multi-componential. There will be improvements in general knowledge about task mechanisms or training material, and there will be improvements in general or specific processes of
cognitive functions triggered by exercise. Practice effect is a typical manifestation of better knowledge. Also learning new strategies mainly targets knowledge of the task rather than enhancing processes. The difference between these two mechanisms is in their direction, as extended practice is bottom-up while applying a strategy is top-down (Zelinski & Reyes, 2009). Lövdén and colleagues (2010) do not speak about near/far transfer (an exclusively behavioural classification) but distinguish between transfer of knowledge and transfer of processes (identifying the neural implementation of transfer). The two classifications are not interchangeable, but in all likelihood far transfer is due to a transfer of processes. Based on this, process modification and far transfer represent the prototypical scenario of plasticity. This does not imply that knowledge-based changes are not due to plasticity: neuroplastic responses may take place equally but will be limited. However, neuroplasticity never refers to improvements due to flexibility or simple fluctuations of performance (Noack, Lovden, Schmiedek, & Lindenberger, 2009).

CA is supposed to trigger a response in the nervous system quantitatively comparable to the response elicited by PA. However, qualitatively-wise, the effects of CA and PA seem to differ. Cardiovascular exercise causes a neural response that likely does not depend on the activity and that is probably homogeneous in a population. There is variability due to different intensity or duration of the training, but the qualitative impact on the brain (and on cognitive domains) should remain the same. On the other hand CA can be modulated and specifically directed to part of the brain and address selected cognitive domains.

For this reason, the following review will focus on part of this vast literature. To begin with, the first section will deal with the behavioural evidence of plasticity that is far transfer in the absence of neurodegenerative conditions. Moreover, the focus will be mainly on older adults, in order to move on easily to the diagnoses of MCI and AD.

### 2.2.2 HIGHLIGHTS ON FAR TRANSFER IN HEALTHY ADULTS

Transfer of processes is possible because the same brain regions are required for processing both training and testing material (Persson & Reuter-Lorenz, 2008), but has been exclusively studied in behavioural research. Reports of far transfer are scarce in the literature, but the potential of these findings must not be underestimated, as it is evidence of a degree of

---

9 This paper was retracted by the authors because of technical flaws in the procedure section discovered afterwards (Persson & Reuter-Lorenz, 2011).
possible neuroplasticity in healthy and abnormal ageing. The most prominent paper was published by Jaeggi et al. (2008). In this study young participants were trained daily for 25 minutes with a spatial n-back task. Those who trained for 17 days or more displayed a significant improvement in a task similar to (but harder than) the Raven Matrices. No gains were registered for those who trained for 8-12 days. The authors explained these results as processes that are shared between the two tasks. Other seminal results are those by Dahlin and coworkers (2008). In this study a form of training based on exercises of working memory with specific focus on updating of various material improved performance in an n-back task and in a verbal recall task after 15 sessions distributed over 5 weeks. These effects, however, were only visible in the younger subgroup. Chein & Morrison (2010) trained young undergraduates for four weeks with complex working memory tasks and found transfer of gains in reading comprehension skills and in a Stroop task, but only in the subgroup of people who actually showed an improvement in training material. Far transfer was reported also by Karbach & Kray (2009), who trained children, young and old adults either with a single-task or dual-task executive control programme. Those in the “switching”-training group obtained a higher improvement in a Stroop task, in verbal and spatial working memory, and in logical reasoning. Importantly, age group showed no interaction with any other variable, suggesting that far transfer is possible even at an old age. A 5-week training of working memory based on a commercially available software (Cogmed QM) resulted in improvement in working memory and an executive test (Paced Auditory Serial Addition Task (PASAT)), suggesting the presence of far transfer at age 60-70 (Brehmer, Westerberg, & Backman, 2012). Importantly, in this paper the control group were not a no-contact group but participants that engaged in the same exercises but maintaining a low-level challenge. Independent high-functioning octuagenarians were selectively recruited and trained by Buschkuehl and colleagues (2008) with computerised working memory material. Training lasted for 12 weeks, with two 45-minute weekly sessions, and the control group exercised very mildly on a cycling ergometer. Transfer to a visual recall task was registered through paired t tests (but not replicated by MANOVA). Other studies failed to find any evidence of far transfer but found evidence of near transfer for old adults (Bherer, Kramer, & Peterson, 2008; Bherer et al., 2005; Li et al., 2008). This suggests that, even if evidence of successful plasticity is not constantly replicated by all studies, at least general chances for cognitive improvement are possible even at an old age.

Most of the training material here reviewed consisted of one or more computer-based tasks. Computerised cognitive tasks have recently been implemented into specific software and videogames. When computer games are merely based on the transposition of “lab” cognitive tasks, they share the same properties, with no or little differences. For example the “Brain
Training*-like series are console-versions of paper-and-pencil training material, with extra dynamic components (e.g. stimuli that move on the display). Despite the commercial claims, a large study showed that computerised stimulation resembling that of famous marketed devices did not live up to expectations. Three ten-minute sessions of training per week were administered for 6 weeks to over 11000 healthy adults and did not even lead to near transfer (Owen et al., 2010). This suggests that technology does not represent a convenient method to subdivide interventions (technology-based vs. technology-free). It is instead more important to clarify which computerised interventions are based on principles of reserve, plasticity, and transfer, and what programmes are instead unspecific or driven by commercial rather than scientific interests.

When games used are actual videogames in the "classical" sense of the word (for example "shooters" or "business simulation games"), they represent an interesting source of training, because it is harder to speculate that any resulting cognitive benefit is due to simple practice or near-transfer. "Vintage" games could not take advantage from today's technology, and the cognitive benefits were limited (Zelinski & Reyes, 2009). New-generation games seem more promising. Basak and colleagues (2008) found that older adults playing a strategic videogame ("Rise of the Nation: Gold Edition") for 8 weeks displayed immediate improvement in executive functions. The team of Green & Bavelier (2008, 2012; 2012) have focused on action video games, reporting improvements for young adults in switching skills, attention and learning. Two weeks of "War of Warcraft" resulted in a trend of improvement in the Stroop performance (Whitlock, McLaughlin, & Allaire, 2012), the hectic "Need for Speed" triggered changes in parameters of structural connectivity (Sagi et al., 2012), and young adults trained for 12 weeks with the "evergreen" Tetris (transposed in 3D) transferred their new skills to tasks measuring spatial abilities, which, according to the authors did not resemble the training material (Terlecki, Newcombe, & Little, 2008). Although videogames are mainly designed for recreational purposes, the possible benefits on old people's cognition could be catalysed by fun aspects (Zelinski & Reyes, 2009). A much more structurally-complex cognitive and emotional response triggered by "proper" videogames could even overwhelm the efficacy of commercialised software whose main goal should be to keep cognition fit: if this were true, it would be paradoxical. However, a review has concluded that there is no evidence for quantitative differences in effectiveness between the two types of stimulation (Kueider, Parisi, Gross, & Rebok, 2012), and moreover, ad hoc software is suitable for being structured and modelled according to requirements, whereas this is not feasible with videogames. The videogames approach, defined by Park et al. (2007) as "non-traditional", does not specifically
aim to enhance knowledge or processes, but taps cognitive aspects that belong to the sphere of leisure activities.

2.2.3 COGNITIVE ACTIVITIES, BRAIN AND COGNITION IN OBSERVATIONAL DESIGNS

Parallel to the literature on PA-questionnaires and risk of developing cognitive decline, there is a branch of research that has investigated cognitive decline as a function of reports of engagement in a list of leisure activities. Comparable strengths and weaknesses can be identified in these studies. Again the large number of participants boosts statistical power in these studies, and again methodological shortcomings can be identified with the most evident one being represented by the variability in the lists of activities used to assess CA. These studies usually include activities such as reading, playing board games, playing musical instruments, dancing and playing cards. In addition some teams chose to include extra activities at their discretion such as going to a museum (Wilson et al., 2010), engaging in group discussion (Hall et al., 2009), being involved in organisations and clubs (Singh-Manoux, Richards, & Marmot, 2003), travelling and karaoke singing (Dodge et al., 2008), painting and drawing (Wang, Karp, Winblad, & Fratiglioni, 2002). As noticed, some of these activities are not exquisitely cognitive (e.g. participation in religious ceremonies, or doing non-exerting leisure physical activity). Nonetheless the pattern of findings appears to be quite homogeneous.

The use of NINCDS-ADRDA criteria has revealed that women (Crowe, Andel, Pedersen, Johansson, & Gatz, 2003) and individuals (Scarmeas, Levy, Tang, Manly, & Stern, 2001; Wilson et al., 2002) who engage in CA more frequently have reduced odds to develop AD at a later age. With comparable methodologies, CA has been reported to be protective as well against dementia (Verghese et al., 2003; Wang et al., 2002) and amnestic MCI (Verghese et al., 2006). Negative associations (of positive meaning) were also found in the longitudinal decline of cognitive functions at a baseline age older than 55 y.o. and with a test-retest interval of 1-to-5 years (Dodge et al., 2008; Leung et al., 2010; Niti, Yap, Kua, Tan, & Ng, 2008; Wilson et al., 2010). The specific association between CA and memory was studied by other teams. A cross-sectional association was found for semantic and memory skills in middle-late adulthood (Singh-Manouxx et al., 2003). Hall et al. (2009) focused instead on the Buschke Selective-Reminding Test and on an older cohort who later developed dementia (age: 75-85 y.o.), following them up with annual assessments to identify the change-point when memory started to decline. The amount of CA was negatively associated with the position of this change-point
on the timeline. In a subsequent paper this change-point was also reported to be significantly slowed down by time spent engaging in simple crossword puzzles (Pillai et al., 2011).

While confirming this general line of findings with their data collection, Iwasa and colleagues (2012) speculated that the association between cognitive hobbies and risk of dementia might actually be the result of an opposite relation of causation. The first symptom of a subtle ongoing pathology could be the gradual withdrawal from cognitive-stimulating activities. Indeed, engagement in everyday activities is reduced in AD (Friedland et al., 2001). To understand this, elderly were recruited by Aartsen and her team (2002) and their cognitive assessment and questionnaire were inputted in a structural equation model, in order to test separate hypotheses regarding the direction of causation. Results support an association but do not appear to support any of the two causal links. A limited effect was found for speed of processing, which suggests that faster individuals prefer more demanding CA (thus, a reversed causation). Based on this, it is possible to conclude that an association has been consistently reported between stimulating leisure activities and cognitive competence, but different experimental designs are needed to understand the nature of such a link.

Other problems associated with these studies are described. First, the use of questionnaires is an indirect and scientifically not manipulatable method to process a variable. Second when composite scores were calculated, they were often derived from activities that also have a social or a PA component, impossible to control for. Third, variability can be detected in procedures of CA-encoding and inferential statistics (Ghisletta, Bickel, & Lovden, 2006). Fourth, weakness is derived from the heterogeneous effect of CA previously described; while aerobic activities are “mono-directional” in the sense that they result in one single picture of neurocognitive enhancement (regardless of the type of activity chosen), cognitive activities appear to trigger an heterogeneous effect. For this reason variation in activities is supposed to play an extra role, in addition to frequency, duration, or intensity (Leung et al., 2010) and window of exposure along the life-time line (Coley et al., 2008).

Some studies used an observational approach to investigate the association between reported engagements in CA and brain volume. A study in which both measures were collected at the same timepoint found a positive association for total amount of GM and WM (Hafsteinsdottir et al., 2012). Longitudinal studies with a 3-year interval between questionnaire and structural scans found no association with neither global (Foubert-Samier et al., 2012; Valenzuela, Sachdev, Wen, Chen, & Brodaty, 2008) nor regional (Foubert-Samier et al., 2012) GM and WM, and a positive association with change of hippocampal volume (Valenzuela et al., 2008). One last study detected no correlation between PiB staining and concurrent amount of CA, but a
significant negative association when past rates of CA were inputted in the analysis (Landau et al., 2012). This indicates that longitudinal designs are preferred, as change scores seem to have more predictive power than “stand-alone” scores.

In conclusion, results coming from observational studies seem to lead to a homogeneous pattern of findings. However, due to the methodological weaknesses listed above, the findings should be taken with extreme caution. Some degree of control on the aforementioned aspects is needed, and control is only possible though the direct manipulation of the activity, which is the main feature of randomised trials.

2.2.4 COGNITIVE ACTIVITY AND COGNITION IN INTERVENTION TRIALS

2.2.4.1 INTERVENTIONS IN HEALTHY INDIVIDUALS

2.2.4.1.1 INTERVENTION STUDIES AND BEHAVIOURAL VARIABLES

Section 2.2.2 focused on aspects related to practice effect and transfer; there is another group of studies in which protocols of CA were tested in relation to cognitive responses from a wider perspective, without a specific highlight on a single task, but with a general intent to improve cognition in healthy elders. The main reason to test protocols of such ambitious nature is driven by the need for non-pharmacological interventions for neurological conditions and by the need of theoretical models to gain a better comprehension of the neural system.

Ageing is associated with neural and behavioural events that are partially similar to those triggered by pathological states, yet remaining in the domain of normality. Trying to contrast healthy-ageing processes is not a priority. This consideration is strictly linked to the lexicon normally used to identify intervention based on cognitive exercises. There is still confusion in the terminology in this field, between the concepts of cognitive stimulation, training, or rehabilitation (Steinerman, 2010) and there is a fine line between interventions with a clinical/theoretical purpose and protocols with a “cosmetic” (Chatterjee, 2004) and commercial objective. Most of the software available to the big public belongs to this latter category.

Papp, Walsh & Snyder (2009) reviewed and analysed 10 disparate trials carried out on healthy samples; unfortunately the nature of the interventions was too variable, ranging from piano lessons to group discussion, to cognitive-motor mixed training. Unless criteria permit the inclusion of a homogeneous set of papers, meta-analyses on CA interventions are not informative, and critical reviews are instead needed. Large cohorts were recruited and trained
with various kinds of CA, based on the use of strategies (Ball et al., 2002), material based on
principles of neural plasticity (Mahncke et al., 2006; Smith et al., 2009), or commercialised
"brain-training” videogames (Ackerman, Kanfer, & Calderwood, 2010; McDougall & House,
2012; Nouchi et al., 2012). This literature is reviewed in Section 5.1.3.1.

Generally speaking, experimental evidence appears less convincing than observational results.
Once again, it seems legit to expect a certain degree of variability when trying to interpret the
effects of a set of variables on a purely behavioural domain. Possibly, findings are more
homogeneous when investigating the effect of CA and targeted interventions on structural and
functional variables of the brain.

2.2.4.1.2 INTERVENTION STUDIES AND NEUROIMAGING
The interest in CA effectiveness has been object of interest during the last few years, and only
recently both observational and experimental evidence has been studied in relation to neural
variables; brain research is an ample branch and the methodology has been highly variable.
For this reason most studies available in this field so far are of disparate quality.

2.2.4.1.2.1 BRAIN STRUCTURE
The study of structural changes induced by CA is a relatively new area of research. It has been
shown that it is possible to observe longitudinal changes in the GM of healthy adults following
a period of motor training (Draganski et al., 2004) and for this reason some teams of authors
have started to investigate possible changes in brain structure due to CA. Programmes based
on CA have been found to trigger GM changes (Takeuchi, Taki, Hashizume, et al., 2011;
Takeuchi, Taki, Sassa, et al., 2011), FA increases (Sagi et al., 2012; Takeuchi et al., 2010) and
cortical thickening (Engvig et al., 2010). These findings are reviewed in Section 5.1.2.1, and are
based on studies in which principally young adults were recruited.

2.2.4.1.2.2 BRAIN FUNCTION
The idea of assessing the efficiency of training through the analysis of brain activity would
provide important information that is independent from brain structure. However, the use of a
task in these fMRI paradigms can be criticised for several reasons: first, as reported in Chapter
1, this type of approach is extremely susceptible to small alterations in the cognitive tasks used
(Trachtenberg, Filippini, & Mackay, 2012), hindering the comparison between two studies that
used a similar but not exactly equal methodology. Second, Stern (2009) postulated that brain
activation does not simply depend on the nature of the task, but also on the interaction
between task load and underlying neural efficiency. Inter-individual differences in cognitive
reserve can alter the activation pattern and create a major source of variability. This would be a particular cause for concern when a group of old adults is compared to a group of young adults, but the authors clearly state that variability in neural efficiency is visible even within a group of adults of similar age. Experimental support for this hypothesis was recently collected. Two latent variables were identified, corresponding to global activation patterns expressed differently in young and old groups of individuals, and one of these patterns displayed group differences in its direction, with young participants recruiting more activation as task difficulty increased, and old participants displaying instead reduced activation (Stern et al., 2012). Third, the use of a single task for the investigation of brain function is limiting. When the efficiency of a CA programme is tested in a clinical sample, a comprehensive battery of cognitive tests should be administered to understand what benefits have been triggered in terms of practice and transfer. An overview of improvement in brain function would be impossible, because it would mean testing all cognitive functions inside the scanner. The use of a single task can bring important evidence about a single cognitive aspect, which is most of the times related to the material used during the training. This means that far-transfer would not be investigated. As far-transfer is the “gold-standard” behavioural manifestation of plasticity, a single-task-associated fMRI would only allow the researchers to study partial effects of CA.

Generally speaking, the use of a different, task-free, paradigm would be a higher-order alternative. First, it represents a standardised measure of general brain function; second, there are interpretative frameworks which have proven effective in the description of healthy ageing and the course of AD-pathology.

The study of CA and changes in measures of resting-state brain function is a new topic of research. CA has been found to modify seed-based and whole-brain resting-state, and all the evidence is reviewed in Section 5.1.1.1.

2.2.4.2 INTERVENTIONS IN MCI

Testing the efficiency of CA on patients is qualitatively different than carrying out research in samples of healthy adults: as detailed before, CA based on principles of neuroplasticity ideally do not rely on strategies; asking (or training) individuals to use different strategies to perform a task means asking them to do something different from what the task was originally meant to measure. This is particularly important for cognitive tests with clinical relevance and for neural activity related to task performance in functional neuroimaging. However, when it comes to judging the benefits of any protocol on everyday life of a pathological population, any sort of training may be beneficial. Strategy training could trigger an equally significant neural response, which would be particularly important for diagnostic entities such as MCI and
AD. Strategy training was investigated by various authors, sometimes with success (Belleville et al., 2006; Moro et al., 2012), sometimes not (Londos et al., 2008; Unverzagt et al., 2009). For this reason, even if recent evidence has shown that strategy training does not lead to any benefits with untrained material neither in healthy individuals, nor in MCI (Hampstead et al., 2012), it is necessary to be flexible when categorising and discussing different types of CA in relation to pathology.

There is a set of papers that have not tested the efficacy of CA programmes but rather the applicability of a single mechanism of training, usually with a single task. Face-Name-Association tasks have been piloted in MCI and AD (Clare, Wilson, Carter, Roth, & Hodges, 2002), together with the Errorless Learning paradigm (Clare et al., 2000; Jean, Simard, van Reekum, & Bergeron, 2007; Lubinsky, Rich, & Anderson, 2009; Metzler-Baddeley & Snowden, 2005), cueing techniques such as the Vanishing Cues (Bier et al., 2008; Haslam, Moss, & Hodder, 2010) or the Significant Event Technique (Moore, Sandman, McGrady, & Kesslak, 2001) and the Spaced Retrieval paradigm of recall or recognition (Bier et al., 2008; Lee et al., 2009).

Several papers have described programmes of training aimed at targeting computational processes in MCI. There is large variability in the essence of these programmes. For instance, a team chose to rely on slightly less rigid inclusion criteria, and enrolled participants older than 65 who were free from relevant diseases and under no AD-related medication, providing them with exercises for either 8 or 12 weeks; the programme was based on exercises like working with paper, wood or clay, and more cognitive training like tasks based on proverbs and verbal material. Improvements were registered in the MMSE score, but only for those who displayed an abnormal cognitive level at baseline (Park, Kwon, Seo, Lim, & Song, 2009). Other programmes (reviewed in Section 5.2.3.1) had a more cognitive signature (Barnes et al., 2009; Cipriani, Bianchetti, & Trabucchi, 2006; Gunther, Schafer, Holzner, & Kemmler, 2003; Rozzini et al., 2007; Talassi et al., 2007), but it is not easy to draw a set of conclusions from these studies, as they represent isolated attempts to test the efficacy of CA programmes of diverse characteristics.

A few experimental studies investigated the relation between CA and neural variables (Belleville et al., 2011; Hampstead et al., 2012; Hampstead et al., 2011; Rosen, Sugiura, Kramer, Whitfield-Gabrieli, & Gabrieli, 2011). This evidence is reviewed in Section 5.1.1.1.
2.2.4.3 INTERVENTIONS IN ALZHEIMER’S DISEASE

2.2.4.3.1 TERMINOLOGY
Since AD is a condition related to impaired independence in everyday life, methodology in research for this population has not been extremely strict, and theoretical weaknesses have permeated the literature: CA programmes based on tasks that are not strictly cognitive, the absence of a control group, the impossibility of separating the effect of the training from that of the cholinergic drugs, the impossibility of performing a complete neuropsychological battery, the use of CA programmes individually shaped on each patient. The presence of a control group would ensure that any improvement in a cognitive test are not the mere effect of practice (Clare & Woods, 2004), but unfortunately it is not always possible to satisfy this methodological condition in a clinical setting. For all these reasons objectives should be recalibrated to a lower level of ambition. Cognitive interventions in relation to a diagnosis of AD are a heterogeneous picture. Many different CA subgroups exist, each of them having their own characteristics and targets. Unfortunately there is a bit of confusion in the literature, because sometimes different names are chosen for similar protocols and sometimes the labels are used interchangeably: Farina et al. (2006) distinguish global and specific interventions, where global ones aim to generally re-orientate the person to a context of reality, whereas specific ones aim to target determined cognitive areas. A similar distinction is made by Clare et al. (2003) and Cipriani et al. (2006), who distinguish cognitive rehabilitation and cognitive training, or by Sitzer et al. (2006), who classify activities according to their compensative or restorative purpose. A more detailed classification discriminates cognitive rehabilitation, cognitive training and cognitive stimulation (Clare & Woods, 2004; Vidovich, Lautenschlager, Flicker, Clare, & Almeida, 2009), and one step further is made by those who separate neuropsychological rehabilitation, memory rehabilitation, reality orientation and cognitive stimulation (Yu et al., 2009). From this perspective, a first classification axis can be established on the basis of the therapeutic intention. Interventions whose goal is to stimulate specifically or generally cognitive functioning, and interventions whose aim is to improve as much as possible the functional level of the person. A second axis can be then represented by the specificity of the intervention. A strictly cognitive programme may be composed by a few tasks only (tapping memory, for example) or on the other hand may consist of a global set of different tasks (memory, plus attention, plus language...). The same can be applied to functional level-centred research, where interventions may be multidimensional or they may focus on a single component only (memory rehabilitation for instance).
The use of various labels is not a problem as long as one clarifies their exact meaning. When "cognitive training" is not used in its general sense as a synonym of "cognitive intervention", it specifically means engagement in a determined set of tasks targeting specific cognitive domains (Clare & Woods, 2004). Similarly, when "cognitive stimulation" is not used in its general sense as a synonym of "cognitive intervention", it aims to broadly target the cognitive functioning in a more general way (Clare & Woods, 2004; Eckroth-Bucher & Siberski, 2009). Moreover, it is referred in particular to a context of group, whereas cognitive training focuses more on the individual (Vidovich et al., 2009). "Cognitive rehabilitation" attempts to restore the physical, psychological and social well-being as much as possible (Mimura & Konutsu, 2007); it does not focus on a single domain but tackles the level of disability in a broader way, often encompassing various types, not necessarily cognitive, of stimulation: for example the paper by Kurz and his team (2009) contains "cognitive rehabilitation" in its title, but their programme is based also on relaxation, motor exercises and stress management. Even though the core of the rehabilitation programme may be exclusively cognitive, the main concern is above all on everyday functioning and ADL. Many types of activity may have their role in a non-pharmacological programme for cognitive rehabilitation (Cohen-Mansfield, 2001; Olazaran et al., 2010) and some of them may even sound funny (Rosler et al., 2002), but all attempts share a common point: tapping on remaining cognitive resources and the use of external aids. Compensatory strategies like the Method of the Loci belong to the category of cognitive rehabilitation, rather than training (Craik et al., 2007; Londos et al., 2008).

2.2.4.3.2 INTERVENTION STUDIES AND BEHAVIOURAL VARIABLES
General meta-analyses have reported conflicting results. Cochrane criteria identified 9 studies of cognitive rehabilitation published between 1988 and 2004 fulfilling satisfying methodological severity, but the results yielded no significant benefit associated with CA in samples of AD patients with average MMSE scores ranging from 17 to 24 (Clare et al., 2003). In this analysis no specific attention was given to the presence of concurrent medication treatment and some of the participants were diagnosed with vascular or mixed dementia. The same nine plus additional eight studies were considered by Sitzer and colleagues (2006). They reported beneficial effects of CA especially over learning and executive functions. Unfortunately, individual and group treatments were taken together, and interventions were based on either restorative (training) or compensative (rehabilitation) criteria. Reviews suffer from the same problems. For example Grandmaison & Simard (2003) reviewed only memory training in samples of AD patients who were not on medication. Even if they eliminated one big confounding factor, other shortcomings prevented the authors from drawing strong conclusions (sample size above all, which was always equal or fewer than 11 participants).
There are paradigms that are considered as “cognitive” even if they focus on aspects that are extremely basic. Reality Orientation Therapy (ROT) belongs to this category, as it focuses on the sense of reality of the patient, who may have lost their orientation in space and time. The aim of ROT is reorienting constantly the patient in space and time, providing the person also with some memory information related to themselves. It can be implemented 24 hours a day, but is normally used when AD patients have already seriously declined, and orientation appears to be the only reasonable therapeutic horizon to aim for. Benefits of ROT have been reported in general measures of cognition (Metitieri et al., 2001; Spector, Davies, Woods, & Orrell, 2000), and although CA has recently moved forward to new and fruitful paradigms, ROT may still represent a flanker strategy (Woods, 2002) and has been confirmed to enhance the benefits resulting from Cholinesterase Inhibitor (ChEI) treatment (Onder et al., 2005).

While interventions based on CA were quite homogeneous in samples of MCI and healthy individuals, a big variability is detectable in AD. ROT and cognitive stimulation (based on lexical exercises, famous faces, use of money and current affairs) were administered to a large sample of moderate AD for 7 weeks, triggering improvements in MMSE and ADAS-COG scores (Spector et al., 2003). The ADAS-COG subtests in which the benefit was statistically significant were "commands" and "spoken language", together with the language subscale (Spector, Orrell, & Woods, 2010). Further MMSE improvement was displayed by those who were enrolled in 16-week period of maintenance, compared to those who stopped training at the end of the first part of the trial (Orrell, Spector, Thorgrimsen, & Woods, 2005). Different trials obtained alternate findings: multidimensional cognitive rehabilitation has proven to be significantly effective in a small sample of MCI but not in AD (Kurz et al., 2009), whereas another paper reported maintenance of cognitive levels in mild-to-moderate and severe AD after 2 years of treatment (Yanguas et al., 2006). The team led by Farina (2002) recruited mild AD patients on ChEI treatment and trained them with a programme either based on procedural memory or cognitive abilities that were partly spared. The first group benefitted from stimulation and obtained significantly better scores in the digit cancellation test and a p value approaching significance for the letter fluency task. In a subsequent paper (2006), the team tested both stimulations combined in relation to a “global” training involving recreational activities. Interestingly, the latter group displayed a significant better difference score in phonemic fluency. The interpretation given to these findings highlights the benefit associated with the social sphere of stimulation, an aspect that is rarely involved in CA. In another article, a small sample of mildly-moderately AD demented ex-career soldiers was recruited and assigned to a mock or a real training intervention. Training tasks were based on ROT, perceptual, executive and learning skills. After 10 weeks an improvement in MMSE was reported in the experimental
group (Niu, Tan, Guan, Zhang, & Wang, 2010). Unfortunately, like other studies described above, separate t tests were performed. Furthermore, the authors chose to include a verbal fluency task and a perceptual task which resembled two tests used in common practice (namely the phonemic fluency test and the Poppelreuter-Ghent Overlapping Figures Test). Although it is not advisable to use cognitive tests as training material, in this particular study they were not part of the testing material.

These trials were designed in light of a clear rehabilitating purpose. Often a general measure of cognition has been the only cognitive outcome and a multifaceted (often individualised) training has been chosen, based on benefits of enjoyable activities, social context, and simple everyday activities. To date, the only observational study found that engagement in leisure activities after AD diagnosis is associated with slower rate of decline (Treiber et al., 2011).

In the continuum between training and rehabilitation, there are other studies that can be located closer to the opposite pole. These studies have a bigger theoretical footprint and draw from models of reserve and plasticity. Reading and arithmetic exercises led to improvement in the FAB, while the MMSE remained stable (and got worse for the controls) (Kawashima et al., 2005). Unfortunately, the two groups were analysed separately. Other papers described the use of computerised CA. Fernandez-Calvo et al. (2011) used the Nintendo Wii Big Brain Academy for 12 weeks on a sample of drug-treated mild AD. ADAS-COG scores remained stable unlike in the control groups in which dropped. Tarraga et al. (2006) trained mild AD participants with a multidimensional cognitive stimulation programme for 12 weeks, and reported slight, yet significant drug-exercise synergic benefits on MMSE scores, with a significant maintenance detected by ADAS-COG. A recent publication reported that 3 months of lexical-semantic exercises improved measures of global cognition, language and verbal memory in mild AD patients who were not on ChEI treatment (Jelicic et al., 2012). Benefits survived correction for multiple comparisons; this has been the first intervention designed in light of evidence of decline in semantic memory associated with early-AD stage (Vogel, Gade, Stokholm, & Waldemar, 2005). It is worth pinpointing how the same material (a marketed videogame) was found beneficial in AD (Fernandez-Calvo et al., 2011) but did not exert any improvement in normal cognition (Ackerman et al., 2010): a further sign that the competence of the participants significantly influences the final outcome.

A parallel branch of research has investigated the effects of CA in relation to AD-related medication. In these studies two groups were compared in relation to cognitive scores collected at baseline and after treatment: a group receiving only pharmacological treatment and another group receiving both pharmacological treatment and CA: interestingly, various
types of CA have been tested with this experimental paradigm: combination of drugs and ROT was more beneficial for the ADAS-COG score than drugs alone in a sample of mild-to-moderate AD treated intensively for 3 weeks (Giordano et al., 2010). The ADAS-COG score was also positively affected by 8 weeks (12 hours in total) of cognitive-communication training. A group of AD patients of comparable severity benefitted more from CA and medication together, rather than just medication (Chapman, Weiner, Rackley, Hynan, & Zientz, 2004). A small sample of AD patients were either assigned to a drug-only condition or to a mixed treatment based on medication and CA consisting of communication, learning and verbal-fluency tasks. All patients had already been on medication at the beginning of the study. After 20 sessions (first administered at a weekly rate, then one session every two weeks) a significant group-by-time interaction revealed a positive synergic effect of medication and CA on MMSE scores (Matsuda, 2007). The same pattern of findings was replicated in a shorter intervention, with only a 30-minute session for seven weeks of specific exercises (Matsuda et al., 2010). Finally, a mild pharmacologically-treated AD group obtained improvement in MMSE score and digit span backwards after 5 months of various exercise modalities (ROT, errorless learning, communication and ADL training), compared to the drug-only group (Bottino et al., 2005).

Other studies preferred controlling for the confounding effect of medication by enrolling only participants who were already under pharmacological treatment. In the study by Löwenstein and colleagues (2004), better cognitive scores emerged after 12 weeks for mild AD who trained with exercises of rehabilitation compared to a control group who exercised with recreational activities. Similarly, verbal learning and fluency tests benefitted from a stimulation programme published in Spanish (Cassinello, Mestre, & Fernandez-Ballesteros, 2008).

2.2.4.3.3 INTERVENTION STUDIES AND NEUROIMAGING

Most of the published fMRI studies have investigated the changes triggered by face-name association training on activation required to perform a task of the same nature (van Paasschen, Clare, Woods, & Linden, 2009). Goveas and colleagues (2011) have recently measured the efficacy of a pharmacological treatment with a ChEI in terms of improvement in parameters of resting-state brain activity. Enhanced functional connectivity was reported, with an increase in positive and negative correlations between the hippocampus and a set of widely distributed areas, with parallel improvements in the cognitive subscale of the ADAS battery. As mentioned above, non-pharmacological interventions could represent a valid alternative to medication, and their efficacy could be similarly measured as functional recovery assessed through parameters of resting-state brain activity theoretically inserted within the DMN framework.
In summary, the impact of CA has been extensively studied in relation to behavioural and neuroimaging variables. General lack of theoretical models has emerged. Most of the protocols were built on observational or clinical evidence and no experimental hypotheses or pattern of predictions were put forward beforehand. Cochrane reviews have determined that engagement in cognitive stimulation results in significant improvement in cognitive skills both in absence (Martin, Clare, Altgassen, Cameron, & Zehnder, 2011) or presence (Woods, Aguirre, Spector, & Orrell, 2012) of dementia. However, albeit studies included in Cochrane reviews are all characterised by high-quality methodology, a qualitative discrimination is necessary to detect the aspects that make the programme effective; from this point of view, the most promising findings derive from studies based on principles of plasticity that take into account neural and cognitive evidence.
3 AIMS AND OBJECTIVES

There is experimental evidence demonstrating that AD starts developing in early adulthood. This early stage is phenotypically silent as it does not trigger any remarkable consequence in behaviour. This means that somebody hit by sporadic AD will still live the majority of their life with no subjective or objective cognitive impairment. In comparison, there are other neurodegenerative diseases whose full-blown forms have an earlier age of onset and a more fatal course (i.e. amyotrophic lateral sclerosis). Despite this, AD has high rates of incidence and prevalence, and these are destined to increase during the next decades. For this reason effective preventive and therapeutic measures for AD are urgently needed. Non-pharmacological interventions may be an alternative or complementary approach to ChEI treatment, but at present evidence consists only of sporadic and diverse studies that have not led to any structured pattern of findings.

A major methodological issue has been the failed identification of an operational paradigm of intervention. CA and PA, the two most studied forms of non-pharmacological stimulation, have often been used within no theoretical framework. Recent studies on brain function have shown that AD causes a progressive failure in brain connectivity (Section 1.3.4). The idea of resting-state connectivity as a measure of neural health is more dynamic and hierarchically more informative than a picture of brain structure or task-related brain function. For this reason it could represent a valid looking glass to describe test-retest changes induced by a treatment. At the moment there are a few papers that have investigated the impact of PA and CA as a function of changes in connectivity. CA can be freely moulded as needed to target specific brain regions and would be the best candidate to target specific networks in order to modulate parameters of connectivity. PA cannot be tailored to target specific set of brain areas, and for this reason it offers less flexibility. At present all we know is that a PA treatment aimed at increasing cardiovascular fitness seems to trigger a benefit in the prefrontal cortex and in cognitive functions that tap this region, but no satisfactory explanation has been put forward. In other words there is enough evidence to test CA efficacy using a connectivity framework, but it is still necessary to find a valid explanation for the selective effects of PA on the frontal lobe.

The objective of this study was to add important state of the art to for use of PA and CA, by identifying the most fruitful paradigms of study and interpretation and testing specific hypotheses in relation to a possible implementation in AD prevention and treatment. In
addition, the role of ApoE polymorphism was studied in relation to the effects triggered by PA and CA. In more detail we aimed to:

A. Test the impact of a single session of exercise based on oxidative metabolism on brain function and cognition of a specific sample of completely untrained adults.

It can be inferred that the same physiological mechanisms are involved in the neural response to AE and CE. For this reason AE represents a methodological alternative to the study of CE on brain function and cognition. The only theoretical frame that predicts a regional effect of AE on brain and cognition, and that is supported by neuroimaging and physiological evidence, was only tested on samples of young fit adults. We recruited a sample of adults who had fitness levels similar to a sample of old adults in order to speculate what could be the effect of exercise in a population at risk of developing AD. The findings of this study are reported in Chapter 4, Experiment 1.

B. Test the effect of stress response triggered by exercise on brain function and cognitive performance.

It has been reported that stress hormones have an effect on neural activity, and the target areas are the same influenced by AE. For this reason it is possible that stress response may account for a significant portion of variability of cognitive function measured under condition of metabolic perturbation due to AE. Since stress is also a risk factor for the development of AD, we investigated its role in a single session of exercise. The findings of this study are reported in Chapter 4, Experiment 2.

C. Understand the effect of the ApoE genotype on brain function and cognition during AE.

Differences in brain function have been reported between young carriers and non-carriers of the $\varepsilon_4$ isoform of the ApoE gene. One of the areas where these differences have been observed is the prefrontal cortex, which is also influenced by AE. For this reason it can be hypothesised that adults engaging in AE would have different levels of exercise-dependent perturbation in brain function according to their genotype for the Apoe gene. As this idea could imply that genotype for the ApoE may drive the efficacy of a treatment based on PA, we tested this hypothesis in a sample of healthy young adults. The findings of this study are reported in Chapter 4, Experiment 3.
D. Create a programme of CA based on principles of neuroplasticity, and tailored to restore brain networks that are normally affected in early AD stages.

Specific cognitive exercises were created with the intention to trigger regional connectivity between areas that are susceptible to early AD pathology. A one-month programme of intensive computerised stimulation was administered to a sample of healthy elderly adults. Changes in brain structure, brain function and cognition were investigated. The findings of these studies are reported in Chapter 5, Experiments 4-6.

E. Test the effect of a CA programme on brain and cognition in a sample of MCI individuals.

The same programme of CA was administered to a sample of patients with a potential neurodegenerative process. Changes in cognition, brain structure and brain function were analysed to: understand the level of retained plasticity during the MCI stage; define the quantity and quality of changes that it is possible to induce with a specifically-tailored non-pharmacological stimulation. The findings of these studies are reported in Chapter 5, Experiments 7-9.

F. Test the effect of the ApoE genotype in non-demented adults who engaged in a programme of CA.

A small mixed subsample of individuals who were administered the CA programme carried the $\varepsilon_4$ isoform of the ApoE gene. Measures of their cognition, brain structure and brain function were compared to a subgroup of noncarriers that received the same treatment, matched by age, education and diagnosis. The findings of this study are reported in Chapter 5, Experiment 10.
4 SHORT-TERM TREADMILL EXERCISE, NEUROCOGNITIVE IMPACT, AND THE ROLE OF APOLIPOPROTEIN E AND STRESS RESPONSE

As introduced in Chapter 2, CE is a sequence of AE sessions structured in a week/month-long intervention. Given this connection between the two paradigms it could be assumed that, if CE has a preferential impact on some areas of brain and cognition, AE will trigger an effect in exactly the same domains. The precise mechanisms by which CE influences the brain are still unknown (Foster et al., 2011), but it is believed that induced neuroplasticity gives a significant contribution (as suggested by the neurobiological hypothesis). Vice versa, according to the model of Lövdén and colleagues (2010), as AE paradigms are based on one session only, plasticity must not be implicated. Nonetheless, AE represents a convenient avenue to study the effect of PA on brain and cognition, as it is not as time-demanding as CE and permits to recruit bigger samples. One of the main inconveniences associated with this paradigm is the difficulty to access neural data. Obviously no significant structural changes can be expected after one session of exercise, but, as for brain function, obtaining pure measures is difficult as data are highly affected by exercise-associated movement. For this reason, most of research has focused on cognition only.

In AE paradigms, participants undergo a short bout of exercise, and a cognitive measurement is performed before, during, immediately after, or after the session. Usually, young adults have been studied, and often one cognitive domain only is tested (with the occasional addition of a control task in case of research driven by an initial hypothesis). The purpose of AE is causing a temporary change in physiological variables that could provoke a repercussion on cognitive skills.
4.1 EXPERIMENT 1 - THE IMPACT OF ACUTE EXERCISE ON BRAIN AND COGNITION

4.1.1 INTRODUCTION

At present, the interpretation of the effects of CE does not go beyond the cardiovascular or neurobiological hypothesis. If no neural interpretation is taken into account for AE, the risk is again to review a series of disparate papers with no chance of drawing overall conclusions. Indeed, there have been a considerable number of papers that have in some way tested the impact of AE on cognition, and the literature may look confusing at a first glance. Single sessions of exercise have been found to have an extremely variable impact: positive effects, negative effects, no effect, linear associations, curve associations: all have been reported by different studies (Tomporowski, 2003). Given the disparity in the available evidence it is necessary to identify a series of variables that are reputed to play a crucial role in the apparently intricate relation between AE and cognition (Brisswalter, Collardeau, & Rene, 2002). The first key-variable is the type of exercise, as activities that trigger different metabolic pathways may cause different impact on the brain. Exercise enhancing levels of cardiovascular fitness has been extensively studied in relation to brain and cognition, and for this reason only AE, principally based on the oxidative pathway will be taken into account. A second factor is the cognitive domain tested. CE studies have come to the conclusion that the frontal lobe is the brain area where the main alterations caused by exercise are observed (Colcombe et al., 2006; Kramer & Erickson, 2007). In the frontal lobe are the neural components responsible for executive processes (Miller & Cohen, 2001), and most of the benefits in cognition have been recorded in executive functions (Chapter 2). A second cognitive domain that has been frequently investigated is memory. The reason behind the interest in memory processes lies in the nature of PA programmes. CE is seen as a potential preventive tool against the onset of cognitive decline, and AD represents a major cause of concern in the ageing population. As memory problems are the main feature of prodromal AD, PA has been specifically studied in relation to its possible preventive function, and thus in relation to early-AD memory decline. As the effect on cognition of a single session of exercise is likely to be labile, and as experimental designs have been usually outlined to investigate one cognitive domain only, a methodological choice must be made. It must be decided on which single cognitive function to focus. AE has been also used to perturb a range of other cognitive aspects, such as attention (Sanabria et al., 2011) or learning skills (Winter et al., 2007), but no satisfying body of research exists on these
functions. Executive functions represent the ideal behavioural domain to explore: the reason for this is theoretical. The impact of AE on executive components is supported by both neuroimaging evidence and a strong set of neurocognitive hypotheses. A major problem with literature on AE is that most evidence of effects on cognition was reported in exploratory studies that did not have any clear prediction based on neural evidence.

To date, the only published set of experimental hypotheses predicting regional effects on brain function is the Transient Hypofrontality Theory (THT), first put forward by Dietrich (2003). The THT takes into consideration the whole amount of metabolic resources available in the brain, which is supposed to be finite, and the qualitative impact of a single exercise session of high intensity on these resources. Motor and sensory areas will heavily draw on metabolic availability, and, complementarily (and following the postulate which suggests that different neural systems compete to “pillage” the available resources), other areas will be temporarily inhibited. This inhibition would be proportional to exercise intensity and duration and can be interpreted as a progressive involvement of a hierarchical model depicting cognitive skills not directly necessary to engage in mere physical exercise. This model is a pyramidal representation of brain regions involved in consciousness, in a graphic hierarchy from the top (responsible for highest-order consciousness) to the base (simple arousal) (Dietrich, 2003).

Electrophysiological evidence from animal research has determined that the top of this hierarchy is represented by the prefrontal cortex (Fuster, 2000), and more specifically by its dorsolateral portion (DLPFC) (Dietrich, 2003), which would be the first area to be affected by an exercise session. The DLPFC is normally involved in top-down cognitive processes in which active cognitive control is needed (Miller & Cohen, 2001). As intense exercise continues, more areas of the pyramid become involved and generally the person starts desiring to disengage from the activity before many components of the model are affected. For this reason the DLPFC would be the core of the impact of AE.

Although the THT does not predict any global metabolic or cardiovascular change in the brain, an assumption crucial to the THT is that global rate of cerebral blood flow (gCBF) does not increase during exercise of heavy intensity compared with a resting state (Dietrich, 2009). Research on human physiology has addressed this question. Past methodology determined that no difference exists between gCBF at rest and during exertion (Globus et al., 1983). Recently, however, new and more sophisticated paradigms have been introduced to measure CBF more accurately (Ogoh & Ainslie, 2009), and these methods have found that mild-to-moderate exercise triggers an increase in global brain perfusion (Ogoh & Ainslie, 2009), but intense exercise decreases this level to baseline level (Hellstrom, FischerColbrie, Wahlgren, & Jogestrand, 1996).
This quadratic trend of gCBF in relation to exercise intensity does not coincide with the trend depicting the relation between exercise and global cerebral metabolism: CBF is vital to maintain adequate metabolic levels during exercise because metabolic resources become available through the blood stream; as a mechanism of feedback, it is likely that levels of global metabolism control gCBF. Oxidation is the only metabolic pathway that produces a considerable amount of ATP in the brain, as the output of the lactate and phosphocreatine anaerobic systems is limited (Secher, Seifert, & Van Lieshout, 2008). The global picture of mutual association between oxidative pathway and global metabolism is not normal when exercise intensity is sufficiently heavy: as intensity increases, the cerebral arteries desaturate, inducing hyperventilation and increases in uptake of oxygen, in order to keep metabolism constant even if CBF is reduced (Nielsen et al., 2002). This means that the system tries to keep the brain in standard conditions but this entails a modification of the relation between metabolism and perfusion.

All things considered, it can be concluded that in conditions of intense AE an altered picture of physiological supply sprinkling the brain is visible, but global availability of resources does not increase. This pattern suggests that the physiological assumption at the base of the THT is valid.

Further support for this assumption comes from the sole neuroimaging technique that is unaffected by movement artefacts: Near Infrared Spectroscopy (NIRS). NIRS measures the dynamics of haemoglobin with good temporal resolution, and provides with information about blood oxygenation levels of cortical mantle located close to the skull (Ekkekakis, 2009). Given its non-invasiveness, it has been easily associated to AE paradigms, and specific focus on the frontal lobe confirms the results found in physiological studies. Despite a low spatial resolution and the impossibility of measuring oxygenation patterns in subcortical areas and midline regions in their inferior part, this technique has provided evidence that exercise modifies frontal perfusion in a way that is proportional to session intensity. Mild and moderate intensities provoke an increase in forehead’s oxygenation whereas heavy intensities cause a steep decrease (Bhambhani, Malik, & Mookerjee, 2007; Ekkekakis, 2009). A functional subdivision of the frontal lobe highlighted how this trend is visible in the motor and premotor cortices, but is particularly pronounced in the prefrontal cortex (Subudhi, Miramon, Granger, & Roach, 2009). Evidence of exercise-dependent reduction in prefrontal metabolism has also been suggested by a PET study (Kemppainen et al., 2005), but the measurement of glucose metabolism was not a real-time observation of brain function. Since physiological and neural evidence corroborate the THT and suggest that DLPFC-mediated cognition would be particularly affected by exercise, the Event-Related Potentials results of Kamijo and colleagues...
(2007) represent a useful bridge between neural substrate and cognition. Testing participants after exercise sessions of progressively-increasing intensity, these authors found a quadratic trend in the P3 amplitude associated with performance in an Eriksen Flanker Task. The P3 amplitude is proportional to attentional and executive resources spent in a task (Kida et al., 2004), and is influenced by exercise during an Eriksen Flanker Task (Hillman, Snook, & Jerome, 2003). The picture of findings suggests that there is a direct link between the impact of exercise on DLPFC brain function and a drop in well-defined cognitive skills.

Since activity of the DLPFC is requested during working memory (Curtis & D'Esposito, 2003) and executive functions (Sylvester et al., 2003), tasks tapping these cognitive domains are the appropriate candidates to test the THT. Indeed, as a first attempt, specific negative impact of heavy AE was registered in the performance in the WCST and PASAT, whereas no effect was visible for two tasks not requiring any cognitive control (Dietrich & Sparling, 2004). Similarly, Del Giorno and colleagues (2010) reported diminished performance in the WCST and the Contingent Continuous Performance Task (an attentional task with an executive component) during intense cycling. This pattern of convincing results is paralleled by other studies having found apparently conflicting evidence. Davranche and colleagues carried out two experiments testing the effects of AE on the performance in two executive tasks: a Simon task and an Eriksen Flanker task. The Simon Effect was negatively affected by exercise (Davranche & McMorris, 2009) but no impact was reported in the Eriksen Flanker Effect (Davranche, Hall, & McMorris, 2009). This may appear surprising, considering that both tasks are based on management and resolution of interference (Zhang, Zhang, & Kornblum, 1999). This apparent confusion can be solved by adding a third explanatory variable to the model (in addition to type of activity and cognitive domain studied) that can disentangle the conflicting findings and illustrate the cognitive effects of the THT: task difficulty. Classic Stroop and Simon tasks require the participant to inhibit a feature of the stimulus that is processed automatically, and to respond according to a different feature, normally processed less automatically. The Eriksen Flanker task used by Davranche and colleagues (2009) was a coloured version of the classic paradigm, in which the participants had to respond following the colour of the central circle, ignoring the two coloured circles placed on both sides of the target. As pointed out by the authors themselves, the feature to ignore in their Eriksen Flanker Task did not belong to the target, but was located externally, possibly making the task easier. A similar conclusion may be drawn from the absence of effect in a task that does not only tap executive skills but also a substantial memory component (Coles & Tomporowski, 2008). Reasonably, any metabolic unbalance can only be visible if a consistent amount of cognitive control is needed.
A fourth important factor is exercise intensity. The THT predicts prefrontal inhibition for heavy intensity regimes. The top of the quadratic relations described earlier would suggest that moderate intensities facilitate executive processes. Pontifex et al. (2009) found that moderate aerobic activity benefitted performance in a working-memory-based modified Sternberg Task. Sibley et al. (2006) instructed their participants to exercise at a self-paced intensity, finding improved performance in the Stroop interference test. It has been shown that the self-adjusting exercise conditions results in participants selecting a more comfortable and less physiologically challenging intensity (Lander, Butterly, & Edwards, 2009), which should ideally be moderate. Finally, a NirS investigation studying the neurocorrelates of a Stroop Task during moderate cycling found increased prefrontal perfusion and improved performance (Yanagisawa et al., 2010).

A fifth factor is the time of task-administration. It is likely that human metabolism can recover quickly from an acute non-pathological unbalance, and for this reason the THT would be valid only if the task is administered while the metabolic perturbation is still on-going. Indeed, Netz et al. (2009) found an improvement in cognitive flexibility (another “prefrontal” skill), when measured 40 minutes after the end of the exercise. Del Giorno et al. (2010) extended the original THT idea, finding an effect of aerobic exercise on prefrontal cognition not only during but also at the immediate end of the session.

A sixth factor, sometimes neglected by reviews, is represented by baseline fitness level of the sample. The same experimental design might lead to different results according to the participants’ level of fitness. Trained adults are able to monitor their physical conditions better than untrained adults, and they are also able to predict some important information of various nature relative to the immediately upcoming minutes of their performance. Untrained participants may feel less confident regarding these aspects, and this could have an important impact on their exercise performance and, as a consequence, on the cognitive tasks. Untrained adults also display a bigger variability in physiological response to exercise (Hotting et al., 2012). In all the aforementioned studies the participants were reported to be physically trained or, more generally, described as capable of undergoing an exercise session, without any precise details about their fitness levels. On the other hand, Netz et al. (2009) specifically reported a differential impact of AE depending on baseline fitness levels. Hogervorst et al. (1996) found an improvement in word-colour Stroop performance immediately after relatively intense exercise, in a procedure that can be compared to the one by Del Giorno et al. (2010). In this case the discrepancy can be detected in sample choice. In the older paper the participants were endurance athletes while in the latter they were simply reported as having “an above-average fitness levels”. The ability to recover from exercise is one of the skills
athletes are more trained in (Du et al., 2005), so perhaps the shape of the inverted U curve previously described might display a later decline in such a sample. From these data it appears that prefrontal-dependent cognitive skills may be particularly impaired after exercise in those people who do not fall in the category of athletes or, generally speaking, fit people.

Comprehending the picture of AE and cognition and hypothesizing a common pathway of influence between the effects of AE and CE on brain and cognition can sensibly increase the pattern of overall converging evidence towards the involvement of the frontal lobe and executive function as areas of exercise effect. There is however a major methodological concern that could be raised. The THT, together with most of the evidence gathered in AE paradigms, is based on studies in which young participants have been recruited, whereas the effects of CE have been studied in samples of elders. Old adults are at more immediate risk of developing AD, and would receive a clinical benefit from PA. Findings obtained from young samples could be partially incompatible with them, and it is not perfectly clear whether generalising the results is legitimate. However, it should be noted that much of animal research has been carried out on exercise, and results have been very often transposed to humans with caution, but also with a degree of confidence judged as fair. It is likely that the physiological impact of PA, if not completely, is at least partially species-independent. For this reason generalisation of results from an age-range to another age-range within the same species is a theoretical leap that does not seem as big as the between-species leap. It might be possible that an age-dependent impact of AE is actually spurious, masking a pure effect of fitness. In all likelihood young participants are fitter in average than old people, and generalisation of results to a different fitness-group has already been questioned (Hillman et al., 2003). Electrophysiological but not behavioural differences were detected after AE between older and younger adults in an Eriksen Flanker task (Kamijo et al., 2009).

As described in Chapter 1, age-related differences in brain function are expected; for this reason AE should trigger qualitatively similar effects on brain metabolism, but calibrated with the age-dependent differences in brain function. However, if fitness level is more important than age, it would be worth investigating the impact of AE in a group of unfit adults, rather than a group of old adults. Since AE paradigms have been extensively used with highly-fit samples, it would be informative to test the THT for the first time in a sample of low-fitness adults, regardless of age. We wanted to test this, and we also decided to give particular relevance in our experimental design to at least a measure of subjective response to exercise. The reason behind this choice lies in the main difference between our study and the studies of AE published in the literature: exceptions aside, AE paradigms have been tested in samples of endurance athletes, fit participants who could comfortably sustain an exercise session, or
individuals whose fitness level was not clearly described. As a specific test on untrained participants has never been performed, it is possible that the psychological reaction to AE displays a substantially big variability in this subgroup. For this reason a measure of subjective response to exercise may be a helpful variable.

4.1.2 METHODS

4.1.2.1 PARTICIPANTS

Seventy-two healthy adults were recruited among the university population, and external volunteers were also included. Demographic characteristics of the participants are reported in Table 1. Age did not differ between genders, $F_{[1, 70]} = 1.126; p>0.05$. Participants were initially briefed about the basics of the experiment to make sure they had an overall picture of the procedure. No specific inclusion criteria were used. The list of exclusion criteria was as follows: age older than 40, chronic disease that could affect muscular, endocrine, cardiorespiratory or neural response, acute illness, significant musculoskeletal injury, use of medication that could significantly prevent them from performing an AE session, an oral disease or problem that could alter the measurements taken from saliva, a drug treatment that could potentially interfere with the metabolic systems involved in exercise, neural activity and cognition, systolic or diastolic arterial blood pressure respectively higher than 160 and 100mmHg. Further methodological exclusion criteria were insufficient knowledge of the English language and a condition of abnormal vision or colour-blindness. All participants signed a form declaring that they had not been physically active during the latest 6 months at least, and they also completed and signed a medical questionnaire to detect any potential threat or medical condition that could potentially interfere with the experimental manipulation.

Ethical approval was obtained from the Ethical Committee of Hull University Psychology Department. Informed written consent was obtained from all participants.

Table 4.1 Demographic characteristics of the sample$^{10}$.

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Age Range</th>
<th>Gender (M/F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>20.76</td>
<td>18-38 y.o</td>
<td>25/47</td>
</tr>
</tbody>
</table>

$^{10}$ All tables in this paper display age and education level expressed as arithmetical mean and standard error.
4.1.2.2 MATERIAL

4.1.2.2.1 ACUTE EXERCISE MANIPULATION

An AE paradigm based on oxidative metabolism was chosen, and treadmill sessions were selected to manipulate the aerobic workout. A Cateye Fitness EC-T200 treadmill was placed in a quiet room for this purpose. Effort was made to keep the environment comfortable for treadmill exercise, and environmental conditions were kept as constant as possible to avoid any possible interference. A cardboard panel was cut out and taped to the treadmill to prevent participants from seeing the main display during the treadmill sessions, and allow at the same time the experimenter to read the information about speed, clock and HR, and manipulate running speed at his discretion by pressing the speed-up and speed-down buttons. The panel was specifically designed to leave the 'Stop' buttons perfectly within the participants’ reach.

A Polar® T31 wireless HR transmitter was tightly worn at chest level by all participants for the duration of the experiment, to allow the experimenter to monitor their HR online. Ultrasonic transmission gel was spread on the transmitter at the beginning of the experimental session, whenever a HR signal test was not reputed satisfactory.

The speed of the treadmill was determined from the response of cardiac output. The "220-minus-age" formula was applied to estimate the maximum HR of each participant. A 70% of this value was chosen to modulate the speed of the treadmill for the duration of the experimental condition. The minimum speed of the treadmill (1.0 km/h) was the constant velocity of the control condition.

In the experimental condition the participant started walking on the treadmill and 1-2 minutes were spent to allow the experimenter to modify the speed and induce a heartbeat at 70% of the estimated full capacity. Participants exercised for 10 minutes at fixed cardiac response, and treadmill speed was constantly monitored in order to keep the HR on the calculated value (± a 5-beat interval). After ten minutes of exercise the participant was asked to sit in front of the computer and start the first task straight away. Two sessions were included in the experimental condition to allow for the administration of 2 cognitive tasks.

In the control condition a 1-2 minute mock warm up preceded the 10-minute walk. The experimenter pressed the slow-down button from time to time (with no effects on the speed) to mimic the speed modulation during the experimental condition. Two sessions were included in the control condition.
4.1.2.2 COGNITIVE TASKS
To test the effects of AE on neural-dependent cognition two cognitive tasks were selected. The logic behind the choice of the two tasks lies in the picture predicted by the THT and the areas normally tapped by each of the two chosen tasks: one of the tasks was chosen to rely specifically on prefrontal cortex to test the THT; the other one was chosen among those that do not require any prefrontal processing and represents a methodological control to avoid any misinterpretation of a general overall effect that might be independent from the THT. Computerised versions of a Stroop and Posner tasks were chosen to assess cognitive performance under effort condition.

4.1.2.2.1 STROOP COLOUR-WORD INTERFERENCE TASK
The classic Stroop paradigm normally consists of three parts: the first section is a written (usually on paper) list of words indicating colours structured in columns and printed with black ink, which the participant is asked to read as quickly as possible; the second part is a sequence of coloured dots presented in the same way, and the task consists of naming the colours as rapidly as possible; the third part is a list of words indicating colours, but this time the meaning is different from the colour of the ink every single word is printed in. In this condition participants are requested to ignore the meaning of the word (a feature that is normally processed with high automaticity) and to respond according to the colour of the ink of the stimulus (a less automatically processed feature). We chose a computerised version of this task, consisting of two sets of colour words subdivided in a congruent and incongruent condition. The Stroop colour-word interference indicates the average difference expressed in milliseconds needed to respond to incongruent trials (mismatch between features) in relation to congruent trials (match between features). This interference, also called Stroop Effect, is associated with the differential contribution of areas within the prefrontal cortex (Peterson et al., 2002; Zysset, Schroeter, Neumann, & von Cramon, 2007).

4.1.2.2.2 POSNER SPATIAL CUEING TASK
Posner paradigms are normally implemented in computerised tasks: participants are requested to dislocate their attention from a temporary spatial cue and relocate it to a stimulus popping up on the screen after the cue has disappeared. The Posner Effect normally refers to the difference expressed in milliseconds between the average RT for incongruent trials (the stimulus does not appear in the position suggested by the spatial cue) and congruent trials (the stimulus does appear in the position suggested by the cue). Under normal conditions, brain activity requested by this task activates the temporo-parietal junction and the superior parietal lobe, without any involvement of prefrontal areas (Vossel, Weidner, Thiel,
& Fink, 2009). The proportion of congruent and incongruent trials was kept close to 50% so that participants would not form expectations and involve a frontal computation (Vossel et al., 2009). Stimulus onset asynchrony was variable across trials to prevent the participants from absorbing an internal rhythm of response.

4.1.2.2.3 TASK ADMINISTRATION AND PROCESSING
The Stroop and Posner tasks were selected because they both permit the analysis of RT and error trials, they both have congruent and incongruent trials, and a difference score (effect) can be calculated in both cases. The number of trials was manipulated in order to obtain two tasks of similar duration (about 6 minutes).

The cognitive tasks were implemented in E-Prime (Psychology Software Tools, inc.), and administered using an Iiyama® Prolite E2607WS 26" monitor and a Microsoft® Natural Keyboard. The D, F, J and K keys were covered with a coloured label (in order: red, blue, black and yellow) for the purposes of the Stroop Task. The testing position with the computer consisted of a table and a chair and was placed a couple of steps away from the treadmill, to allow the participants to step off the treadmill, sit down and start the tasks immediately. The monitor was at approximately 80cm of distance from the participant’s eyes.

RT and accuracy were registered for each of the two tasks in each condition; mistakes were separated from correct trials, and correct trials were split according to congruency; all responses of 50ms or less were discarded and a first average was computed, together with a +/- 2 standard-deviation interval. All measurements that fell out of this interval were discarded and an adjusted mean was recalculated to be included in the analysis. RT for error trials were not analysed.

4.1.2.2.3 SUBJECTIVE MEASURES OF EXERTION
Levels of perceived exertion were measured with the Rate of Perceived Extentation (RPE) scale (Borg, 1982). The RPE is a simple score from zero to ten that describes the psychological perception of exertion caused by exercise. The score range is meant to cover the whole spectrum of response, from 0 (“Nothing at all”), to 4 (“Somewhat strong”), to 10 (“Very, very strong (almost max)”).

4.1.3 PROCEDURE

Each participant was tested on two separate days: in the first day one of the two conditions was completed together with the initial screening and questionnaires; on the second day the remaining condition was completed. Both meetings were scheduled at the same time of the
day to avoid fluctuations in performance and hormonal levels. The order of conditions was counterbalanced as well as the order of the tasks. On day one the participant was explained the various steps of the procedure and inclusion/exclusion criteria were examined to determine the suitability for participation. After that the first session started, immediately followed by the first task. No recover was conceded, and the second session started immediately, with the same modalities (Figure 4.1). On the second day the remaining condition was completed with the same modalities as the first condition. The RPE questionnaire was administered after each treadmill session.

**Fig. 4.1 Procedure of the behavioural part of the study**

Data were analysed with the Statistical Package for the Social Sciences (SPSS-PASW) software (v.18). Two-tailed hypotheses were tested, and a significant $p$ value was set at 0.05. The Kolmogorov-Smirnov test was chosen as test of normality, and the Mauchly’s test as test of sphericity. Whenever the assumption of sphericity was not met, the degrees of freedom were adjusted through the Huynh-Feldt correction. As age was used as a covariate in most of the analyses, it was assumed beforehand that it was independent from the variability observed between conditions and among the levels of the between-subjects variables. The assumption of homogeneity of slopes was tested for the main parametric analysis only, by modelling a univariate ANOVA, with the main outcome as dependent variables, the predictor as independent variable and age as covariate: main effects were asked for both predictor and
covariate separately and an interaction between the two variables was also added, to make sure it was not significant.

### 4.1.4 RESULTS

RPE scores (average of the two sessions) were almost always zero in the control condition (mean=0.87, SEM\(^{11}\)=0.16), whereas in the experimental condition this value was close to "somewhat hard" (mean=4.02 (SEM=0.21)). This difference was significant: \(t_{71}=13.85, p<0.001\). As RPE scores for the control condition were not informative, they were discarded from further analyses. The two RPE scores of the two experimental sessions were divided according to the cognitive task that followed (Figure 4.2). The two values were highly correlated (Pearson’s \(r=0.835, p<0.001\), and did not differ statistically (\(t_{71}=-1.298, p>0.05\)). Nonetheless, it was chosen to keep them separated (RPE\(_{\text{Stroop}}\) and RPE\(_{\text{Posner}}\)).

![Fig. 4.2 Distribution of RPE scores in the experimental condition for each of the two tasks.](image)

Both RPE\(_{\text{Stroop}}\) and RPE\(_{\text{Posner}}\) were converted into a categorical variable after division into three tertiles (cRPE\(_{\text{Stroop}}\) and cRPE\(_{\text{Posner}}\)). The first level (scores from 0-“Nothing at all” to 3-“Moderate”) was coded as "Not hard exertion"; the second level (4-“Somewhat hard” scores) was coded as “Moderately hard exertion”; the third level (5-“Hard” to 10-“Very, very Hard (Almost max)”) was coded as “Hard exertion”. The choice to keep the two RPE scores

---

\(^{11}\) SEM: Standard Error of the Mean
separated was further supported by the fact that there was within-subjects disagreement between $cRPE_{\text{Stroop}}$ and $cRPE_{\text{Posner}}$ levels for 26 participants.

Fig. 4.3 Distribution of error trials in both tasks: no significant differences were reported between conditions with regards to incongruent trials. Congruent error trials were not analysed.

Data regarding the Posner task were discarded for one participant because they did not understand test instructions. As the number of errors made during congruent and incongruent trials for both tasks did not distribute normally, specific focus was reserved for incongruent trials only. Nonparametric tests revealed no difference between conditions for the Stroop ($\text{Wilcoxon Signed Ranks Test}= -1.41, p>0.05$) and the Posner ($\text{Wilcoxon Signed Ranks Test}= -0.24, p>0.05$) tasks (Figure 4.3). Explorative 2X2 (condition-by-congruence) parametric ANOVAs confirmed these findings.

We computed the Posner and the Stroop Effect and, as the intention of the study was to compare different conditions within the same sample, the assumption of normality must be met in the distribution of the differences between conditions (Field, 2009): no violation was found. The differences were also used in the test for the assumption of homogeneity of slopes,
following the procedure previously described, imputing age as covariate and cRPE as predictor. No violation was found for the Stroop task, but the assumption was not met in the Posner task: a visual inspection of the graph suggested that the observation associated to the oldest participant of the sample (age=38 y.o.) was responsible for the inhomogeneity; for this reason that observation was removed from subsequent analyses of the Posner task. The assumption was met after the observation was removed.

2X2 ANCOVAs were run on the Effects including cRPE scores and adjusting for age (Figure 4.4). For the Posner Effect, a significant effect of condition was found ($F_{1,66}=5.74, p<0.05, \eta_p^2=0.08$), superseded by a significant condition-by-cRPE interaction ($F_{2,66}=3.53, p<0.05, \eta_p^2=0.097$). This interaction was further inspected splitting it along the axis of the between-subjects variable. Paired t tests were run for each of the cRPE levels, revealing no significant differences between conditions for any of the levels (respectively: $t_{30}=1.237, p>0.05; t_{20}=1.487, p>0.05; t_{17}=-1.433, p>0.05$). Another attempt was made splitting the interaction along the axis of the within-subjects variable. Two separate one-way ANOVAs were run with cRPE as an independent factor, but again no $p$ value was significant (experimental condition: $F_{2,69}=0.57, p>0.05$; control condition: $F_{2,69}=1.03, p>0.05$). This may indicate a significant variability in the general data, rather than a specific effect focusing between two of the levels of the variables.

For the Stroop Effect, a two-way interaction was found ($F_{2,68}=3.19, p<0.05, \eta_p^2=0.086$). As a further investigation paired t tests were run for each of the cRPE levels, revealing no significant differences between conditions when the experimental session was perceived as “Not hard” ($t_{30}=-0.18, p>0.05$) and “Moderately hard” ($t_{20}=-0.65, p>0.05$). A significant difference was instead present when the judgement was “Hard” ($t_{19}=2.77, p<0.05$). Interestingly, including gender as a further between-subjects factor revealed that this pattern of results was only evident for females, who also displayed a significantly smaller SE during the experimental condition when the session was perceived as “not hard” ($F_{2,44}=6.53, p<0.005; t_{19}=-2.32, p<0.05$; data not shown).
Fig. 4.4 RT (correct trials only): on average Posner trials were faster than Stroop trials. No significant difference was found in the Posner task within any level of the cRPE, suggesting that general variability in the data was responsible for the interaction (4.4a). A condition-by-cRPE interaction was also found in the Stroop task, but subsequent analyses revealed that a significant difference between conditions was evident only in those who perceived the experimental session as “Hard” (4.4b).

* p<0.05
4.1.5 DISCUSSION

We wanted to test the THT for the first time in a sample of markedly untrained participants. Although we tried to keep the stimulation constant at moderate/moderately-high levels, we were conscious that unfit adults could have responded to the sessions in an extremely variable way. For this reason we included a psychological judgement of the experimental manipulation as a between-subjects factor in our analysis, and ignored any possible uninformative main effect of condition. A condition-by-cRPE interaction revealed that the Stroop Effect was significantly larger in the experimental condition only when AE was perceived as hard. No similar impact was recorded for the remaining levels of the categorical variable, suggesting that the detrimental impact of AE on prefrontal-dependent cognition can be only observed if exercise is heavy enough. No significant effects were concurrently found for a Posner Task not depending on prefrontal areas, because the significant interaction was accompanied by no significant between-levels or between-conditions difference. Furthermore, a simple look at the graph (Figure 4.4a) suggests that when the stimulation is perceives as “Hard”, there is a trend for facilitation. This pattern of findings supports the THT and suggests that the response of prefrontal metabolism of untrained participants is similar to that of fit adults and athletes. We found significant effects with a stimulation estimated to induce a 70% intense workout. It can be concluded that the intensity necessary to cause a significant neural change is slightly milder in an untrained group, compared to a fit group. It is likely that old people (regardless of the presence of a neurodegenerative disorder) are physically unfit, and their prefrontal cortex may be affected by PA of medium intensity. A programme of CE would reflect a repetitive induction of prefrontal hypometabolism. It is possible that any benefits in frontal areas and/or executive functions/working memory observed after long-term exercise programmes at old age or in the presence of a MCI/AD diagnosis are actually connected to a training-induced optimisation of prefrontal metabolism. In other words, thanks to a momentary reduction of available resources, prefrontal areas would receive a progressively beneficial training. In fact, the recovery from a condition of hypofrontality has been connected to enhanced availability of resources (Yanagisawa et al., 2010) and cognitive performance (Netz et al., 2009). The fact that this pattern was only found in females is open to various interpretations; gender differences have been found elsewhere (Porcelli et al., 2008), and some authors have argued that this discrepancy is due to sex-differences in the HPA axis (Lautenschlager, Cox, & Kurz, 2010).

The idea that CE is beneficial to the brain because it induces an optimisation of prefrontal metabolism is a new perspective that does not challenge the cardiovascular and neurobiological hypotheses, but instead offers a more appropriate explanation to the evidence
collected so far. Both the classic hypotheses describe important physiological mechanisms that should affect the brain globally. The translation of the THT principles on CE paradigms would instead predict a regional change in brain function and cognition, which goes in the same direction as the evidence suggested by the literature.

The analysis of the full group revealed no beneficial effect of light or moderate AE neither on prefrontal-dependent nor on prefrontal-independent cognitive tasks. Conversely, a beneficial effect of AE on prefrontal function was found in the subgroup of women that perceived the session as “not hard”. This finding is in line with the pattern obtained in NIRS paradigms, as light exercise is supposed to momentarily increase brain function in the prefrontal cortex. Our interpretational view posits that repeated hypofrontality would result into an optimisation of prefrontal metabolism. This, however, does not rule out the possibility that equal benefits might be obtained through a programme of lighter sessions. It is possible that repeated exercise-related hyperfrontality due to AE of weak exertion would induce persistent changes in frontal metabolism at rest. Available evidence in the literature does not support this view (see Sections 2.1.3.1.1 and 2.1.3.1.2), but it might be possible that a benefit might be visible after a longer period of light exercise, especially on the basis of a between-group comparison in which a control group remained sedentary for a similarly longer period. However, a specific test of this hypothesis would be required to provide support for this speculation.
4.2 EXPERIMENT 2 - THE ROLE OF STRESS RESPONSE

4.2.1 INTRODUCTION

As the pattern of exercise-related hypofrontality reflects a specific part of the brain, it is possible that this pattern will be subjected to the effect of any mediator that has a specific regional impact on the same neural substrate. There is evidence that the physiological response triggered by stress hormones is a localised mechanism of action specific for a set of brain areas, and the prefrontal cortex is one of these. “Stress” identifies a pattern of events occurring in the organism as a response to external stimuli that acquire a particular level of importance because they are “novel and/or unpredictable” and the person feels they “do not have control over the situation” (Lupien, Maheu, Tu, Fiocco, & Schramek, 2007). These “stressors” are usually linked to particular situations of evolutionary importance (for example “fight or flight” scenarios, or situations of psychological discomfort), and trigger a sequence of molecular secretions of adaptive role.

Stressors activate a hormonal cascade involving several structures of the body. The first response originates from the hypothalamus, which secretes the Corticotropin Releasing Hormone. This molecule activates the pituitary gland, which, in turn, produces the Adrenocorticotropic Hormone. A major receptorial site for this hormone is located in the adrenal glands, where the final products of this cycle are secreted: catecholamines and glucocorticoids. Altogether, this system is called Hypotalamic-Pituitary-Adreanal (HPA) axis. Catecholamines comprise adrenaline and noradrenaline, two molecules responsible for the involvement of the sympathetic branch of the autonomous nervous system. Cortisol is the main human glucocorticoid, and is the only stress hormone that can cross the blood brain barrier and have direct access to the brain, where it can bind to its receptorial populations. Receptors for cortisol are widespread throughout the whole body (Dedovic, Duchesne, Andrews, Engert, & Pruessner, 2009) but specific brain receptors are located in preferential areas. The idea of a molecule being able to enter the brain in such particular environmental situations suggests that it may elicit some immediate significant effects on brain function (and cognition). It is also likely that the impact of this molecule will be specifically associated to the function receptorial areas are normally responsible for. Animal research was comprehensively reviewed by Lupien et al. (2007), who concluded that two main types of receptors exist, and they display an exclusive degree of affinity with cortisol. Mineralocorticoid Type I receptors display higher affinity and are mainly located in mediotemporal/limbic areas. Glucocorticoid
Type II receptors display less affinity and are instead more scattered in the neocortex. Studies on primates have shown how the prefrontal cortex is particularly rich in this type of receptors.

It is necessary to make a distinction between chronic and Acute Stress (AS). This differentiation is similar to that separating AE and CE: observational studies have investigated the impact of chronic stress on the brain, cognition, cognitive decline and AD (usually using cortisol concentration at awakening, or averaging cortisol levels obtained at different timepoints during the day). AS has been instead used as an experimental manipulation to trigger a physiological response of short duration. Several methodologies have been used to induce AS: for example AE (de Vries, Bernards, de Rooij, & Koppeschaar, 2000; Kemppainen et al., 2005), psychosocial stimulation (Fiocco, Poirier, Joober, Nair, & Lupien, 2008), "thermic discomfort" (Porcelli et al., 2008) or exogenous administration of glucocorticoids (Hsu, Garside, Massey, & McAllister-Williams, 2003). Differences have been reported in the neural systems involved in stress response of different nature (Dedovic et al., 2009), but findings have not yet allowed researchers to fully understand the complexity of the neural mechanisms independently involved, and agree on a standard paradigm of study.

As for the relation with normal ageing and AD, chronic stress is generally considered as a variable that does not have any specific causal role, but worsens the overall condition (Pardon, 2011). CSF (Popp et al., 2009) and plasma (Csernansky et al., 2006) morning levels of cortisol are higher in AD but not in MCI. It has also been shown that stress is associated with aspects of decreased neuroplasticity in AD (Gould & Tanapat, 1999; Rissman, 2009), but it is not clear whether the position on the continuum between the absence and presence of AD-dementia is a determinant factor in the relation between cognition and stress hormones. Indeed cortisol level was found to be inversely correlated with cognitive functioning in healthy old adults in some studies (MacLullich et al., 2005), but other studies found no association (Csernansky et al., 2006). Although chronic stress is considered a risk factor for AD, morning cortisol levels were not a significant predictor of AD onset after 7 years of follow-up in a large cohort of healthy adults (Schrijvers et al., 2011). The overall picture relating chronic stress to AD is still confusing, but the lack of unquestionable evidence does not reject the idea of stress as a risk factor for AD.

The findings obtained in healthy old adults are more interesting. The whole spectrum of cognitive abilities was analysed as a function of all-day and change-from-awakening cortisol levels in a cohort of mid-age adults. Only the all-day levels were associated with measures of cognition, and memory and executive functions were equally affected (Franz et al., 2011). Vice versa (and curiously), only the change-from-awakening score was significantly correlated with
measures of general cognition (Evans et al., 2011) and executive functioning (Evans, Hucklebridge, Loveday, & Clow, 2012). Declarative memory and global cognition were studied by Gerritsen and colleagues (2011). An association was found between single measurements of cortisol and indices of cognition, but no longitudinal effect was reported. All in all memory and executive functions appear to be the cognitive domains most affected by chronic stress exposure, but it is not clear whether this is effectively due to an exclusive impact of stress response on these domains (as suggested by Franz et al. (2011)) or to the fact that these domains are actually those who have been studied most extensively. Moreover, it is not yet clear whether a possible association with a specific cognitive domain is dependent on the presence or absence of AD (Csernansky et al., 2006).

As “physical exercise increases the plasma concentrations of numerous hormones, notably the glucocorticoids” (Filaire, Duche, Lac, & Robert, 1996), a session of AE may cause neural consequences independent from the normal reallocation of resources, and dependent instead on stress response. It is not completely clear whether AS can ultimately improve or impair cognitive performance, but, learning the lesson from AE, a cautious guess would suggest that it may depend on the cognitive skill investigated and to the degree of involvement of the HPA axis. Reviews of studies on rodents suggest that a key role is played by the proportion of occupation displayed by the two subgroups of receptors. Learning skills would benefit the most from an increase in glucocorticoids when Type I occupation is completely saturated and Type II occupation is about 50% (de Kloet, Oitzl, & Joels, 1999). This idea is impossible to test in humans, but it is likely that there is a quantity of stress reputed as beneficial for specific cognitive skills, and other quantities that are instead detrimental.

The effects of acute changes in peripheral cortisol have been studied in relation to particular cognitive skills in young humans. Executive skills and working memory have been sometimes investigated in relation to stress hormones (Franz et al., 2011). This domain might represent a potential behavioural mirror of the pattern of receptor sites in the cerebral tissue. Exogenous cortisol administration resulted in a negative impact on working memory, when injected dose and requested computational load were both high (Lupien, Gillin, & Hauger, 1999). It was also found that glucocorticoid administration caused an increased number of errors in a Stroop task (Hsu et al., 2003) and abnormal decision making in a gambling task (Putman, Antypa, Crysovergi, & van der Does, 2010). Performance on the WCST was analysed as a function of pre-test cortisol by McCormick and colleagues (2007), who found a gender-by-cortisol-level interaction over the number of errors (similar pattern for perseverative and non-perseverative errors); no differences were instead found in a mental rotation task mainly based on parietal activity. Generally speaking, prefrontal-dependent cognitive skills appear to be generally more
influenced by experimental manipulation than declarative memory or vigilance (Lupien et al., 2007). In another study a psychosocial stressor caused a strong hormonal response in the experimental group, as expected, but no concurrent behavioural impact was discovered (Weerda, Muehlhan, Wolf, & Thiel, 2010). This probably indicates that the burden on working memory and executive functions requested by the task has to be demanding enough to differentiate those who underwent AS from controls. This was also argued by Schoofs et al. (2008), who observed that AS did not constantly cause a decrement in performance on the Digit Span task, which was believed not to be particularly demanding as a working memory task. It is not clear to what extent the Digit Span can be considered a measure of working memory when the two modalities (forwards and backwards) are computed together, as in Elzinga & Roelof’s study (2005). Similar doubts emerge from the findings of another study. Porcelli et al (2008) used a measure of immediate recognition as a working-memory task, which, however, does not seem an appropriate choice.

Interpreting the results by Weerda and colleagues (2010), it is likely that cognitive scores are not as sensitive as neural measures in detecting the impact of a mildly-challenging task based on prefrontal processing, as significant stress-associated effects were found in the same study in task-related brain function on the ROIs chosen by the authors. In agreement with this view, structural (Kremen et al., 2010) and functional (Dedovic et al., 2009; Weerda et al., 2010) neuroimaging evidence has supported the idea that AS response has a direct impact on a prefrontal neural substrate. It would be this change that mediates the effects on behaviour. Following this outline functional neuroimaging appears as the gold-standard paradigm to study the impact of an experimental manipulation of stress response. Cognitive measures would be a less appropriate choice, because the relation molecule-cognition is a bigger theoretical leap than the relation molecule-neural variable, as suggested by the interpretative model of Imaging Genetics (Mattay et al., 2008). However, if a strong theoretical model based on neural evidence is implemented, it is possible to minimise the variability associated to the sole use of cognitive variables.

The THT represents an unconfuted model of neural activity associated to AE, a stimulation that has proved to evoke a stress response. In this scenario the use of cognitive tasks appropriately tapping the prefrontal cortex would permit the clarification of the impact of stress response in a context of perturbation of the frontal lobe. We tested this hypothesis by investigating levels of salivary cortisol and cognitive performance in a paradigm of AE. As a variable of secondary importance, we also collected test-retest psychological responses to exercise to investigate any potential difference between psychological and physiological changes due to a session of AE.
4.2.2 METHODS

The same participants were recruited and the same general methodology and procedures were used as described in Sections 4.1.2 and 4.1.3. Additional variables are described herein.

4.2.2.1 MEASUREMENT OF STRESS RESPONSE
Saliva samples were collected for measurement of stress hormones. Although central levels of hormone would be the best theoretical choice, levels of salivary cortisol are considered valid, methodologically suitable, and poorly invasive (Lippi et al., 2009) compared to extraction of blood samples (Hamilton, 1995). Twenty-four hours values of cortisol contained in saliva are highly correlated to blood levels and with the gold-standard body fluid peripheral source, which is urine (Neary, Malbon, & McKenzie, 2002). For this reason we chose the least invasive method, in a test-retest fashion, following the method of Weerda et al. (2010). A passive-drool technique was chosen. The participants were asked to abstain from food and beverages (water excluded) in the 2 hours prior to sample collection, from alcohol assumption in the 12 hours prior to sample collection, and from any dental work in the 48 hours prior to sample collection. The participants were asked to rinse their mouth thoroughly with cold water and to wait at least 10 minutes before the first sample was taken. A plastic tube was provided and they were asked to try and pool some saliva in their mouth, and to passively allow it to fall into the tube. The procedure was repeated until a sufficient amount of saliva was collected. The tube was inspected visually to detect any possible blood contamination, and was stored as rapidly as possible at a temperature of -80°C.

Salivary cortisol was measured by Salimetrics’ enzyme-linked immunosorbent assay ("High Sensitivity Salivary Cortisol Enzyme Immunoassay Kit," 2011) following the procedure described by the manufacturer.\(^\text{12}\)

4.2.2.2 STAI QUESTIONNAIRE
The State-Trait Anxiety Inventory (Spielberger, 1983) is a tool used to measure general individual levels of anxiety expressed as a trait and fluctuations in state anxiety. Both subscales consist of a 20-item list, and for each statement a score from 1 to 4 must be indicated. After proper correction, the overall scores were two numbers between 20 and 80. We chose to include a parameter of psychological response to exercise to compare it with the physiological response observable through hormonal levels and for statistical adjustment purposes.

---

\(^{12}\) The assay was carried out by Dr Rebecca Vince, Department of Sport, Health, and Exercise Sciences, University of Hull.
4.2.3 PROCEDURE

The main procedure was as described in section 4.1.3. Additionally, baseline saliva samples were collected during the first day, after written consent was given. State and Trait subscales of the STAI questionnaire were administered immediately after. Retest measurements of saliva and State anxiety were collected after each of the first-day treadmill sessions. In the second day baseline and retest measures were collected with the same methodology.

Cortisol and State Anxiety levels modified by AE were analysed as a function of condition and timepoint. We then analysed the association between change scores of physiological/psychological variables and the Posner/Stroop Effect. For this purpose only data related to the experimental condition were analysed: change scores were computed and correlation with Posner and Stroop Effects was analysed.

4.2.4 RESULTS

The 6 state-anxiety measurements (sSTAI) distributed normally, therefore a 2X3 condition (experimental and control)-by-timepoint (baseline, retest 1 and retest 2) within-subjects ANOVA was run. Significant effects of condition ($F_{1,71}=7.99$, $p<0.01$, $\eta^2_p=0.101$), and timepoint ($F_{1.36, 96.32}=4.64$, $p<0.05$, $\eta^2_p=0.061$) were found, superseded by a significant interaction between the two variables ($F_{1.45, 102.68}=9.83$, $p<0.001$, $\eta^2_p=0.122$). A gradual drop from baseline was observed in the control condition, whereas levels remained stable in the experimental condition (Figure 4.6). These effects vanished after adjusting for age or trait anxiety.

A subjective "change-from-baseline" value was computed with the following formula, in order to obtain a single value assessing the anxious response triggered in the experimental condition:

$$sSTAI_{Experimental-Baseline} - (sSTAI_{Experimental-Retest 1} + sSTAI_{Experimental-Retest 2})/2$$
Cortisol assay was not possible for 22 of the 432 tubes, and 10 participants were affected. Apart from one participant whose six tubes were not analysed, and another participant missing 5 measurements, the remaining missing values represented a major problem only when a baseline measure was missed (3 participants). When the problem regarded one of the retests, the other retest was used in substitution. Sixty-seven participants were thus included in the analyses. The six subjective values did not distribute normally and for this reason, we calculated an average change-from-baseline score for each of the two conditions and used nonparametric statistics. The comparison of these values was not significant (*Wilcoxon Signed Ranks Test* = -0.14, *p* > 0.05), indicating that no difference in stress response was detectable between the two conditions. We also ran a parametric 2X3 ANOVA, and only an effect of timepoint was found (*F*₁, 77.74 = 15.02, *p* < 0.001, *η_p²* = 0.198), indicating that stress significantly declined in a similar way in both conditions (Figure 4.7). Adjusting for age did not change these results. Gender was not associated with cortisol at any point (all *p* > 0.05). Similarly to STAI scale, “change-from-baseline” levels of cortisol were computed with the following formula:

\[
\text{Cortisol}_{\text{Experimental-Baseline}} - \left(\frac{\text{Cortisol}_{\text{Experimental-Re} \text{test 1}} + \text{Cortisol}_{\text{Experimental-Re} \text{test 2}}}{2}\right)
\]
Fig. 4.7 Concentration of salivary cortisol: no differences were found between conditions.

Table 4.2 Correlation between the Posner and the Stroop Effect, and psychological and physiological response to exercise

<table>
<thead>
<tr>
<th>State Anxiety-Change</th>
<th>Cortisol-Change</th>
<th>Posner Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Correlation</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>72</td>
<td></td>
</tr>
</tbody>
</table>

| Cortisol-Change | Pearson Correlation | 0.177          |
| Sig. (2-tailed) |                   | n.s.           |
| N               | 67                | 67             |

| Posner Effect | Pearson Correlation | -0.066         |
| Sig. (2-tailed) |                   | n.s.           |
| N               | 71                | 66             |

<table>
<thead>
<tr>
<th>State Anxiety-Change</th>
<th>Cortisol-Change</th>
<th>Stroop Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Correlation</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>72</td>
<td></td>
</tr>
</tbody>
</table>

| Cortisol-Change | Pearson Correlation | 0.177          |
| Sig. (2-tailed) |                   | n.s.           |
| N               | 67                | 67             |

| Stroop Effect | Pearson Correlation | -0.155         |
| Sig. (2-tailed) |                   | n.s.           |
| N               | 72                | 67             |

127
Neither the Posner- nor the Stroop Effect in the experimental condition displayed significant correlations with the sSTAI cortisol change-from-baseline scores (Table 4.2). Similarly, no association emerged in the control condition (data not shown).

4.2.5 DISCUSSION

The main problem of our procedure, with regards to AS, is that it did not evoke any stress response at all. AE is supposed to increase cortisol levels (Filaire, Duche, Lac, & Robert, 1996), but in our study the levels did not increase but rather decreased over time. There might be several explanations for this pattern of evidence. At first it might be possible that the measurement was taken too early, and not enough time was given to the body to “complete” the HPA cycle with the secretion of the final hormone. This was speculated by McMorris and colleagues (2009) who found instead a satisfactory explanatory fit of Adrenocorticotropin Hormone on their pattern of findings, and motivated it as a methodological artefact due to a too-early measurement of cortisol. This eventuality does not seem to be the explanation for our results. As our experimental condition was structured in two sessions, the second retest would have been surely enough to indicate a possible increase of cortisol due to the first session completed more than 20 minutes earlier.

A second explanation could be the incompatibility of a passive drool technique compared to other studies that used a different methodology. The use of Salivettes (Shirtcliff, Granger, Schwartz, & Curran, 2001) or Eyespears (Strazdins et al., 2005) may produce slightly different estimates of cortisol serum levels (Poll et al., 2007). Moreover, it was a bit problematic in the participants with a particularly dry mouth, as in these individuals cortisol concentration is higher in saliva (Shigeyama et al., 2008), and this may have represented a further source of variability in our data.

A third aspect can be detected in the baseline levels of internal stress. There is evidence that environmental factors contribute to the baseline concentration of cortisol in saliva (Scheer et al., 2002), and for this reason it could be suspected that the decrease in cortisol concentration reflects an initial stress response due to feeling of anxiety, novelty, or discomfort. The morning concentration of cortisol in adults younger than 40 is 0.32 ± 0.02 pg/ml (Ahn, Lee, Choi, Kwon, & Chun, 2007), which is similar to the mean values obtained in our data collection (≈0.2 pg/ml). This indicates that the participants did not start engaging in exercise under stress.
conditions. We observed a statistical drop in hormonal levels which indicates stability into the normal range.

A fourth possible explanation lies in too-light exercise intensities. de Vries and colleagues (2000) reported hormonal increases only with a 80% VO$_2$max intensity, but no specific information was given by the authors about the fitness level of their participants. It is believed that to obtain a physiological reaction of similar intensity non-athletes would have to exercise at a lower workload than endurance athletes with enhanced cardiopulmonary function (Kashihara, Maruyama, Murota, & Nakahara, 2009). Given this evidence, although our sample was unfit, it remains possible that our sessions were not intense enough.

A fifth explanation may be the lack of a sympathetic response in our sample. There is some evidence that a catecholamine-based response is necessary for HPA axis to impair working memory (Elzinga & Roelofs, 2005). This possibility does not appear satisfying, as there is evidence of prefrontal-related impaired performance after administration of exogenous cortisol (e.g. Putman et al. (2010)), where no concurrent involvement of the autonomous nervous system is expected. A cortisol increase should not be seen as an absolute response to stressors, as proportions of non-responders were reported also with valid paradigms (Elzinga & Roelofs, 2005). A similar study that recorded an attenuation rather than an increase in stress response is that by Qin et al. (2009), who showed some videos to their female participants as part of their stress-induction paradigm. In both groups (violent videos and non-violent videos) a decrease in cortisol levels was found over time. The authors assumed that the variability in stress response could be sufficient to detect stress-dependent effects in working memory and in working-memory brain activation. They did not find any behavioural effect and only a mild pattern of between-group differences in neuroimaging data, interpreted as significantly more attenuated brain activation during the task and a correspondent more attenuated deactivation of DMN areas. As already explained, the interpretation of task-associated deactivations as DMN is theoretically biased (and Qin et al. (2009) did not do so), but referring to those areas as DMN-areas is correct, and suggests an important interpretational frame. Since brain activity is structured in networks the effect of stress on brain function will not be merely expressed as a change in brain activity of receptor-rich areas, but rather as a change in brain activity of the networks receptor-rich areas belong to. This idea has already been successfully tested (Veer et al., 2011) and it is possible that in the future manipulated brain function will be more often interpreted as a function of networks rather than areas, in studies using AS paradigms.

In conclusion, we were not able to induce an AS response and we conclude that the findings reported in Experiment 1 are independent from stress response.
4.3 EXPERIMENT 3 - THE ROLE OF THE APOLIPOPROTEIN E GENOTYPE

4.3.1 INTRODUCTION

The effect of the ApoE genotype was assayed in some of the observational studies investigating the risk of developing cognitive decline in relation to lifetime PA, but in most instances it was just used as an adjusting factor. Reanalysing the same cohort of Rovio et al. (2005) with comparable methodology, Kivipelto and colleagues (2008) analysed the proportion of AD and dementia as a function of genotype and an estimate of past (21 years) engagement in PA. The likelihood of these diagnoses for inactive carriers was significantly higher than that for inactive non-carriers; no effect was reported on active participants, though. Different but somewhat complementary findings can be extracted from the study of Podewils and co-workers (2005), in which the benefit of PA on dementia progression was only displayed in $\varepsilon_4$ carriers while no effect of PA was detected in non-carriers. Similarly, non-institutionalised (no mention about baseline cognitive status) carriers aged 60-to-69 years old displayed an association between reported PA and score on a cognitive screening tool, whereas the association was not significant for non-carriers. Interestingly, the association was never significant for the cohort older than 70 y.o., indicating that maybe age can mediate this relation (Obisesan, Umar, Paluvoi, & Gillum, 2012). No impact of the ApoE genotype on the effect of PA was instead found by other teams (Ravaglia et al., 2008; Taaffe et al., 2008), after stratification for this variable. In brief, it is not possible to draw any satisfying conclusion from the body of observational research.

The presence of the $\varepsilon_4$ allele is a risk factor for developing AD. For this reason it has sometimes been studied in relation to PA among people who were already demented, or in relation to AD biomarkers. The Dutch study that did found no positive effect of a walking programme on cognitive skills of a sample diagnosed with AD, also analysed the possible interaction with the ApoE genotype. No difference was reported between $\varepsilon_4$ carriers and non-carriers (Eggermont, Swaab, et al., 2009). In a study that tested cognitively healthy elderly adults, those with longer PiB retention, higher CSF concentration of p/t-TAU and lower concentration of CSF A$\beta_{42}$ reported to engage significantly less in PA, and this trend was particularly exacerbated among $\varepsilon_4$ carriers (Liang et al., 2010). In a more recent study by the same team, PiB staining and CSF A$\beta_{42}$ were analysed as a function of the $\varepsilon_4$ allele and PA: main effects of both predictors were found on both dependent variables, with an additional interaction only expressed on in vivo
brain Aβ deposition: exercisers with the ε₄ allele were undistinguishable from non-carriers, while physically inactive carriers had remarkably higher levels of PiB staining (Head et al., 2012).

For AE and trials of CE, the effect of ApoE has never been tested. When participants engage in PA and perform no cognitive tasks, the metabolic imbalance predicted by the THT takes place in relation to a condition of neural rest. It has been shown that ApoE influences resting-state brain metabolism, particularly in the prefrontal cortex (Reiman et al., 2001; Reiman et al., 1996; Reiman et al., 2005), and even at a young age (Reiman et al., 2004). Since AE would cause a temporary inhibition in part of the prefrontal cortex, and at the same time young carriers display differences in resting-state brain function of the prefrontal cortex, these two aspects share the same neural substrate and might therefore interact. Moreover, there is evidence that carriers and non-carriers respond differently to a stimulation of the nervous system (Marchant, King, Tabet, & Rusted, 2010), and for this reason the presence of the ε₄ allele might be an important variable in the causal link between AE and DLPFC-dependent cognitive skills.

4.3.2 METHODS

Participants, methodology and procedures were the same described in Sections 4.1.2-4.1.3 and 4.2.2-4.2.3, with the addition of the methods of analysis for genotyping procedures of the ApoE gene.

4.3.2.1 APOLIPOPROTEIN E GENOTYPE
Non-invasive buccal swabs were used to obtain cellular material for DNA assay (Milne et al., 2006). Buccal swab and baseline saliva samples were collected immediately after written consent was given. The participants were asked to close their mouth and produce as much saliva as possible. Tongue, gums and cheeks were gently rubbed with a Sterile Foam Tipped Applicator for approximately two minutes and the cellular product was transferred to an FTA Indicating Card®. The impregnated card was left to dry in a safe room and was then inserted in a FTA Multi-Barrier Pouch® with a small Silica Gel Desiccant bag. The pouch was sealed and stored at room temperature for genetic assay.

Preliminary analyses were run to assess the methodology. Two 2mm discs were extracted from each FTA card using a Harris Punch Mat® and a Harris Micro Punch®, and were prepared for Polymerase Chain Reaction (PCR) following the instructions provided by the manufacturer.
Each disc underwent 5 sequential washes, 3 with 200μl FTA Purification Reagent® and 2 with a ph 8.0 TE-1 Buffer. The discs were then left to dry and stored at -20°C.

To test for the presence of any possible cross-contamination three FTA cards were selected and DNA-free discs were obtained after the cutting procedure was completed for each sample. DNA-impregnated and DNA-free discs were amplified with an upstream primer 5’ ACT GAC CCC GGT GGC GGA GAC GCG TGC and a downstream primer 5’ TGT TCC ACC AGG GGC CCC AGG CGC TCG CGG, to isolate the sequence internal to the ApoE gene in which the ε2 and ε4 alleles normally differ from the ε3 allele. The complete amplification reaction is described by Ossendorf and Prellwitz (2000). Visualisation of the PCR products on 3% agarose gel confirmed that there had been amplification products only with the three DNA-impregnated discs, and no products were observed for the 3 control discs.

The PCR products were subsequently cut separately with two restriction enzymes: Afl-III and Hae-II (Zivelin et al., 1997) and visualised on polyacrylamide gel in a UV illuminator.

4.3.3 PROCEDURE

Buccal samples were collected immediately after informed consent was given. The rest of the procedure was as in Experiments 1-2.

To facilitate the interpretation of the analyses, we computed an Effect Increment score for both tasks, by subtracting the effect found in the control condition to the effect found in the experimental condition. Positive scores of this variable indicated the extent to which the effect was larger in the experimental condition. Negative scores indicated instead a larger effect in the control condition.

We compared cognitive performance following AE of carriers and non-carriers via univariate ANOVAs, by inputting the Effect Increment as dependent variable and genotype as predictor. Given the size of the sample we chose not to include cRPE in order not to lose too much statistical power.

As the experimental manipulation did not result in any stress response, we did not investigate the impact of the ApoE genotype on this variable.

---

13 Genetic assays were carried out by Dr. Jennifer Waby, Department of Biological Sciences, University of Hull, and Dr. Charlotte E. Dyer, Division of Cancer, Postgraduate Medical Institute, University of Hull.
4.3.4 RESULTS

The genotype was available for 70 participants. According to epidemiological data (sample size: 734 individuals), the proportion of $\varepsilon_2$, $\varepsilon_3$, and $\varepsilon_4$ alleles in UK is respectively 0.089, 0.767 and 0.144 ($p_{\varepsilon_2}$+$p_{\varepsilon_3}$+$p_{\varepsilon_4}$=1) (Corbo & Scacchi, 1999). Since Hardy-Weinberg equilibrium’s formula for the ApoE gene is $1=p_{\varepsilon_2}^2+2p_{\varepsilon_2\varepsilon_3}+p_{\varepsilon_3}^2+2p_{\varepsilon_2\varepsilon_4}+2p_{\varepsilon_3\varepsilon_4}+p_{\varepsilon_4}^2$, (considering just the main alleles), a 70-person sample extracted from a population showing the aforementioned allelic proportions would result in the distribution reported in Table 4.3, compared to ours.

Table 4.3 Proportion of ApoE genotypes according to normative data and in our sample.

<table>
<thead>
<tr>
<th>GENOTYPE</th>
<th>Corbo &amp; Scacchi (1999)</th>
<th>OUR DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_2\varepsilon_2$</td>
<td>0.08</td>
<td>3</td>
</tr>
<tr>
<td>$\varepsilon_2\varepsilon_3$</td>
<td>2.24</td>
<td>1</td>
</tr>
<tr>
<td>$\varepsilon_3\varepsilon_3$</td>
<td>15.86</td>
<td>13</td>
</tr>
<tr>
<td>$\varepsilon_2\varepsilon_4$</td>
<td>0.25</td>
<td>0</td>
</tr>
<tr>
<td>$\varepsilon_3\varepsilon_4$</td>
<td>3.34</td>
<td>5</td>
</tr>
<tr>
<td>$\varepsilon_4\varepsilon_4$</td>
<td>0.18</td>
<td>0</td>
</tr>
<tr>
<td>TOT</td>
<td>22</td>
<td>22</td>
</tr>
</tbody>
</table>

Pearson’s $\chi^2$ was not significant ($p>0.05$) indicating that our data did not differ from the expected proportions. We excluded from the analyses all $\varepsilon_2$ carriers to compare $\varepsilon_3\varepsilon_3$ (n=35) to $\varepsilon_3\varepsilon_4$ adults (n=23).

Both the Posner and the Stroop Effect Increment distributed normally in our sample, while errors rates did not. For this reason we limited our analysis to the RT.

No effect of genotype was found for both the Posner and the Stroop Effect Increment (Figure 4.8). Including the $\varepsilon_2\varepsilon_3$ participants to the group of non-carriers and the $\varepsilon_3\varepsilon_4$ and $\varepsilon_4\varepsilon_4$ participants to the group of carriers did not change the results. The inspection of Figure 4.8 suggests a trend for the Stroop Task, as there is an increment for non-carriers (as predicted by the THT) and a decrement for carriers. As post-hoc analyses in Experiment 1 revealed that findings were only valid for females, we reran the analysis limited to this subgroup (22 non-carriers and 16 carriers). The Posner Effect Increment was again unaffected by genotype, whereas a significant effect was found for the Stroop Effect Increment ($F_{1,36}=4.40; p<0.05; \eta_p^2=0.11$; Figure 4.9). Adjusting for age did not change the outcome ($F_{1,35}=4.27; p<0.05$;
$\eta_p^2=0.11$). No effect of genotype was found in the subgroup of males (n=20; 13 non-carriers and 7 carriers) for any of the tasks.

**Fig. 4.8** Effect of ApoE genotype on cognitive performance in the full sample (n=58).

**Fig. 4.9** Effect of ApoE genotype on cognitive performance in the subgroup of women (n=38).
4.3.5 DISCUSSION

The $\varepsilon_4$ allele is a risk factor for AD. The onset of the disease is normally set in the late adulthood, but accumulating evidence has suggested that the ApoE $\varepsilon_4$ isoform has a significant impact on brain function even at a young age. However, so far there has been no evidence suggesting that differences in cognition may exist at young age between carriers and non-carriers. The main reason for this lack of evidence can be twofold: 1) despite the functional discrepancies, there is no difference in cognition between carriers and non-carriers; 2) the functional discrepancies reflect a cognitive difference so subtle that under normal conditions it is not possible to detect. AE might represent a gateway for these differences to become evident. As suggested by the results by Filippini et al (2009), the Stroop task may be considered an ApoE-dependent task, whereas the Posner Task an ApoE-independent task. We did not find any condition-by-genotype interaction in the Posner task, and only a trend was detected in the Stroop task. We decided to investigate further this trend by reanalysing only the subgroup of women, as in Experiment 1 we found gender differences in response to AE. We believe that women are generally less prone to exercise and may be considered a “purer” group of untrained participants. Support to this idea is given by the evidence of gender differences in baseline levels of cardiopulmonary fitness (Herdy & Uhlendorf, 2011). It is also possible that a deeper explanation may account for this gender-dependent effect: the genetic profile for the ApoE gene influences the levels of total cholesterol, which in turn can be considered as an estimate of the overall cardiovascular status. Presence of the $\varepsilon_4$ allele (Hubacek, Lanska, Skodova, Adamkova, & Poledne, 2008) and male gender (Kolovou, Damaskos, Anagnostopoulou, & Cokkinos, 2009) are independently associated to higher levels of total hematic cholesterol. It is possible that baseline levels of fitness of completely unfit adults can be accounted for by genotype more than gender. In this way, a specific focus on the female subgroup would allow significant differences to emerge. Other explanations are also possible, related to other classes of metabolic variables.

The findings indicate that under acute metabolic perturbation in which less resources are available to the prefrontal cortex, genotype-dependent differences emerge even at a young age, mirroring the findings on brain function reported by Filippini et al. (2009). The subsequent detailed analysis of this interaction did not reveal any significant pattern, not allowing us to draw more detailed conclusions. The visual inspection of the findings (Figure 4.8) indicates that the Stroop Effect was smaller during exertion in carriers, and this appears counterintuitive, as the THT suggests exactly the opposite. These data could easily fit the model of antagonistic pleiotropy introduced in Section 1.4.2.4: the $\varepsilon_4$-dependent protein could be associated to
positive neuroplastic mechanisms in response to external detrimental factors. The $\varepsilon_4$ isoform could play a positive role at a young age, but being then detrimental in advanced adulthood. To better understand this point, it is helpful to describe another mechanism that is supposed to share similar mechanisms: Lupien and colleagues (2007) reviewed older studies describing the relation between hippocampal pruning and brain function: a positive correlation appears to exist at an old age, whereas the correlation is negative at a young age. Mechanisms of considerable synaptic pruning would determine a smaller but better connected young hippocampus, by removing unnecessary synapses and thus optimising the overall connectivity. In the elderly synaptic redundancy is not as evident as in young people and for this reason the same amount of pruning would be detrimental. The $\varepsilon_4$ could be linked to processes of a similar rationale, down-regulating parameters in excess in early adulthood and down-regulating parameters not in excess any longer in later adulthood.

As PA is considered protective against neurodegeneration, the impact of the $\varepsilon_4$ allele may be observed in modulating the benefits of a programme of PA. It is either possible that carriers, displaying a picture of reduced neuroplasticity, would benefit from PA less than non-carriers; alternatively, carriers could display more benefit precisely because of their lower baseline levels, as the range of possible improvement would be larger. These findings suggest that the $\varepsilon_4$ allele could mediate the effects of PA in a trial of CE. As larger differences between carriers and non-carriers have been found in brain function and cognition of elderly adults, it is possible that this mediation would be particularly exacerbated in the aging populations. No trials have investigated this eventuality, and it is not yet clear in which direction the interaction would be apparent.

Variability in our sample must have played an important role, as it is conceivable that the physiological response to exercise is quite variable in a sample of unfit participants. For this reason, cRPE scores described in Experiment 1 could be an important variable in this design. It is possible that the variability accounted for by ApoE genotype masks the effect of a cRPE-genotype interaction. Unfortunately, our sample was too small to include a further 3-level categorical variable, as the number of observations was not sufficiently large for each combination of the two variables. For instance, in the $\varepsilon_4$=carrier-cRPE=3 combination, there were only 4 observations, and these were massively variable, ranging between increments of -8 to 139 seconds. This indicates that a larger sample is needed to test the interaction between the two variables. Alternatively, the hypothesis could be tested in a sample of athletes displaying a homogeneous physiological response to exercise.
4.4 GENERAL DISCUSSION

We tested three exploratory hypotheses using a paradigm of AE. We found support for the THT for the first time in a sample of untrained adults, and we found no effect of stress hormones altered by exercise on performance in the chosen cognitive tasks. Furthermore, a selective effect of the Apoe $\varepsilon_4$ allele was found on prefrontal function.

One main shortcoming to report in our design is the lack of a strong measure of exercise intensity. An objective measure like VO$_2$ max would have permitted the manipulation of treadmill power at the proper speed, tailored on each participant. Unfortunately, it was not possible to run a VO$_2$ max test to exhaustion, and for this reason we chose to use an indirect measure, assuming that the relation between respiratory output and HR is linear (Kashihara et al., 2009). The “220-minus-age” equation has been criticised by Tanaka et al. (2001) for not being an accurate estimate of maximum HR; these authors collected data from a remarkably large sample of adults of various age and calculated a different, more precise formula: 208–(0.7Xage). In our study 53 out of 72 adults were aged between 19 and 21. Applying the classic “220-minus-age” formula for a 19, a 20 and a 21 years old adult would result into an estimated maximum HR of respectively 201, 200 and 199 per-minute beats. Using the formula suggested by Tanaka et al. (2001) these values would be 195, 194 and 193. Considering that an interval of at least 5 beats per-minute is needed when it comes to adjusting the speed of the treadmill manually, we believe that the difference of output between the two equations is negligible.

A second important aspect to mention is the fact that participants provided their RPE score after the completion of the cognitive task. It is possible that their judgement of session intensity could have been modified implicitly by subjective response to performance in the cognitive task: in other words, a performance in the Stroop task perceived as bad might have altered the perception of the exertion. This is a reasonable point, but it is impossible to estimate the Posner or Stroop Effect. It is much easier to be emotionally affected by the number of errors committed, and we found no relation between the number of mistakes and the RPE. For this reason we believe that this issue is not a methodological shortcoming.

We suggested that a theoretical bridge exists between physiological mechanisms associated to AE and CE. Using the same logic we interpreted our findings regarding the ApoE genotype as a possible mediation played by the gene on the effectiveness of a PA programme on brain function. Again, using the same logic but this time focusing on AE, it can be argued that even the effects of a single session of exercise could be studied as a function of networks. The THT itself is based on the relationship between concurrent computations (Dietrich, 2003) and has
been recently expanded with further theoretical constructs referable to other brain structures (the nuclei of the reticular formation), which are involved in a functional circuit to account for the effects of AE on brain function (Dietrich & Audiffren, 2011). Other authors have hypothesised a concurrent emotional response to exercise, and one of the proposed mechanisms was a mutual negative association between activity of the prefrontal cortex and the amygdala during exercise (Ekkekakis, 2009). These examples suggest that brain function is structured in networks even during exercise. Based on this it would be interesting to investigate what happens to the DMN during exercise. Neural artefacts due to movement are an important methodological obstacle in the study of brain function, but perhaps one or more movement-related network could be statistically detected using procedures like independent component analysis. This statistical method identifies a parsimonious number of brain networks (components) that account significantly well for variability observed in the resting state data. For example Rosazza and colleagues (2012) asked for 20 components and identified the DMN in one of them. Interestingly, a second component described by them covered the outline of the ventricles and was extended to the CSF only; the authors interpreted it as the synchronous movement of the CSF induced by the cardiac beat. This is obviously no pattern of activity and it can be easily left out of the interpretation of the results. Following the same logic it would be possible to combine AE and functional connectivity at rest to detect and isolate specific networks activated by movement, and study those activated/deactivated by the pure effect of exercise.

The interpretation of these results is particularly relevant as stimulation for AD. The cardiovascular and neurobiological hypotheses predict a global pattern of benefits for the brain in its entirety, but all the evidence that highlight the frontal lobe as the main computational areas were the positive influence is observed adds important interpretational bullet-points. An enhanced functionality of frontal areas would be beneficial for people at risk, MCI or even full-blown AD even if the frontal lobe does not represent the initial area where the pathology generally originates from. Prefrontal structures have always been addressed as areas responsible of compensatory reallocation of resources (Chapter 1) in a condition where temporal and parietal areas are directly affected in the early stages. In all likelihood the nature of the mechanisms involved in the optimisation of brain function would rely on neuroplastic availability. As the patient progresses along the axis of the pathology, their capacity for neuroplasticity is slowly depleted. For this reason Foster et al. (2011) argued that the pre-symptomatic and prodromal phases would be the stages receiving most of the benefits of physical exercise. It is possible that many other mechanisms are involved in the beneficial impact of CE. An interesting example is that described by Veer and colleagues (2011), who
investigated the recovery phase from a stressful stimulation. They found that this stage was associated with an increased connectivity between the amygdala and midline structures. As the synchronicity between these two areas belongs to that portion of functional disconnection described by the posterior component of the DMN (Buckner et al., 2008), when AS is triggered by exercise, the recovery would add a further regional benefit to the picture, and neuroplasticity-driven improvement in brain function would be extended also to posterior regions of the brain instead of being defined as global or being confined to the frontal lobe only. Veer et al. (2011) used a seed-based approach, and focused on recovery. Extending a similar methodology to different aspects of exercise, as previously suggested, could provide new and important insights.

Schoofs et al. (2008) argued that the choice of the task plays an important role in the measure of the effects of stress on working memory and executive functions. In the past decades this neuropsychological domain was often described as a unitary component. The Supervisory Attention System (Shallice, 1982) and the Central Executive System of the Working Memory (Baddeley & DellaSala, 1996) were seminal models picturing a single executive component, supported by neuroimaging evidence (Desposito et al., 1995). More recently, Miyake and colleagues (2000) carried out a principal component analysis to find out whether a model with multiple executive variables represented a better fit of the data than a single predictor: they were able to identify three statistically independent executive components: updating, inhibition and switching. Considering the WCST as a measure tapping set-switching exclusively (Miyake et al., 2000), the PASAT as a task clearly based on updating, and the Stroop and Simon tasks as centred on inhibition (Zhang et al., 1999), all three independent component have been reported as affected by AE (Davranche & McMorris, 2009; Dietrich & Sparling, 2004). It can be concluded that the THT does not predict a temporary deficit in determined cognitive components, but assumes instead that a brain area is transiently down-regulated and, as a consequence, any task that needs the involvement of those areas will be affected, regardless of the independence of the executive sub-component. For similar reasons, any effect of stress should be measured in relation to a neural region rather than one or more executive components.

We found that prefrontal-dependent condition was transiently disrupted in females only. It was suggested that this could be due to differences in stress response to exercise (Lautenschlager et al., 2010). But we did not find any gender differences in cortisol levels at baseline or retest. In an attempt to interpret these results through the simplest possible explanation, we believe that females are generally less used to exercise and for this reason the THT was particularly detectable in this subgroup. For this reason we did not analyse the
potential interaction between gender and stress response in the analysis of cognitive performance. This is also the interpretation we gave of the finding indicating that differences existed between female carriers and non-carriers of the $\varepsilon_4$ allele in the performance in the Stroop task. We found a positive increment of the Stroop Effect in the subgroup of female non-carriers, as predicted by the THT, and a decrement in the subgroup of carriers. We prefer not to speculate as to the possible explanation for this effect, but we believe this is evidence that allelic differences exists in neurocognitive response to exercise between untrained carriers and non-carriers. This is the first evidence suggesting a differential response to PA according to the presence of the $\varepsilon_4$ allele. It is therefore possible that long-term exercise programmes may exert a differential response which is dependent on the ApoE genotype.
5 INTENSIVE COGNITIVE STIMULATION AS NON-PHARMACOLOGICAL TREATMENT TO RESTORE REGIONAL CONNECTIVITY

The study of cognitive stimulation in the complete AD spectrum is a wide field of research driven by both clinical urgency and need for a strict methodology. Motivated by these two contrasting drives, many trials have been carried out, unfortunately with an overall scarce proportion of success. One of the main characteristics of this literature is the immense variability of methods, material, and procedure. On the contrary, the theoretical background is a common point for all studies, as AD is associated with a well-known picture of biomarkers, which, despite the numerous inconsistencies listed in Chapter 1, still represents the backbone of any therapeutic attempt. The main problem for the lack of a unitary successful paradigm can be located in the interpretational interface chosen to observe the pathology and at the same time judge test-retest benefits of CA: the choice has almost always been a paradigm based on cognitive functions. At the moment, a crucial contribution to a diagnosis of probable AD is given by a reported objective decline in neuropsychological performance (Dubois et al., 2007). This means that, with no cognitive impairments whatsoever, it is not possible to diagnose AD using the classic guidelines. As neuropsychological tests are a key diagnostic tool for the detection of potential presence of prodromal AD in the absence of dementia (Venneri, Turnbull, & DellaSala, 1996), it appears normal to use a paradigm based on cognition for the creation and evaluation of a CA programme. For example, declarative memory is severely impaired in AD, and is one of the earliest functions to decline. Based on this, CA programmes based on memory exercises have been tested as a potential therapeutic tool (Grandmaison & Simard, 2003). Individualised protocols have been designed to target more advanced stages of the pathology (Clare et al., 2002), often in association with cueing techniques (e.g. Vanishing cues (Haslam et al., 2010)); strategy training has been instead explored in mild AD, MCI and at-risk individuals (Jobe et al., 2001; Unverzagt et al., 2009). In the light of a cognition-disrupting decline, these programmes are theoretically flawless, but have been criticised because they are not based on principles of neuroplasticity (Lovden et al., 2010). There is more research based on cognition but with an eye on plasticity, in which transfer effects have been sought and sometimes found (Li et al., 2011), but with no convincing and consistent pattern of findings. In sum, paradigms of CA created from evidence of cognitive deficits in AD have led to sporadic success. Moreover, as mentioned before, choosing the paradigm is not only about designing the stimulation, it is also about assessing its efficiency. The sole use of cognitive tests
as test-retest measures is probably not sensitive enough for a series of reasons. First, the raw score obtained in some cognitive tests is an integer belonging to a determined range, often with a skewed distribution. This prevents the variable from meeting the assumptions required to run parametric analyses. Second, floor and ceiling effects cannot be avoided. Third, if a programme is effective and induces a small but consistent improvement in cognition, this improvement would be recorded as a small increment in the test's raw score: to be detected by a statistical analysis, this increment would have to be observed in a sufficiently big sample: in fact, maintaining standard values of Type I error and power (respectively 0.05 and 0.8) a within-sample $F$ test with one group and two measurements (assessing test-retest change) will display differences in detection of the effect size as a function of sample size: a $\eta_p^2$ of 0.2 can be detected with 10 participants only, but a $\eta_p^2$ of 0.08 would need at least 25 participants (assuming a moderate correlation between the two measures). As an improvement on cognitive tests is likely to be very small (e.g. an average +0.5 in the raw score in the Raven Progressive Matrices Test), this difference could not emerge statistically in a small sample. Fourth, since multiple testing is normally used to assess many cognitive domains, the correction for multiple comparisons would be a sort of statistical guillotine neutralising most of the significant differences.
5.1 A NETWORK-BASED PROGRAMME OF COGNITIVE TRAINING ON A SAMPLE OF AGEING ADULTS

5.1.1 EXPERIMENT 4 - EFFECTS ON BRAIN FUNCTION

5.1.1.1 INTRODUCTION
The picture of cognitive skills is vital for diagnosing AD, and therefore has been the main framework for creating CA programmes. However, these have not led to any therapeutic certainty. Moreover, changes in cognitive measures are not the best choice for assessing CA effectiveness. It appears therefore necessary to find another interpretational frame to look at AD for creating CA programmes and testing their effectiveness. It would be ideal if the same paradigm could be utilised for both moments. The analysis of cognitive skills is very accurate at placing the person on the clinical stage axis. A similarly-good biomarker would be needed: for example, regional accumulation of hyperphosphorylated TAU protein in the brain is highly correlated to cognitive deficits (Braak & Braak, 1991). As a working methodology it could be possible to identify the areas affected by TAU-related lesion, define what cognitive skills are mediated by those areas and create a CA programme based on this frame. However, using central TAU as a measure of treatment effectiveness is impossible for obvious reasons. A better picture would be represented by measures of brain structure: MRI techniques would be extremely sensitive to test-retest differences. In addition cognitive material could be designed as a function of regional atrophy and cognitive domains related to the areas where a structural decline is observed. However, it is likely that no major differences exist between exercises targeting cognitive difficulties and exercises targeting atrophy-dependent cognitive difficulties. The idea that emerges is that training material based on a view of regional damage in AD (as well as that based on pure cognitive deficit) is generally not satisfying, and that a completely different interpretational framework is required.

AD has also been described as a decline of resting-state brain function (Kogure et al., 2000; Mosconi, 2005), but by using fMRI at rest it is possible to obtain information that goes beyond the simple pattern of areas affected by the pathology. This technique allows the experimenters to identify the functional relation expressed by these areas and how this is hit by AD. A statistical procedure like ICA can identify the latent networks and account for variability in brain function from a hierarchically superior point of view. AD is not seen as a disruption of areas, but as a disruption of their functional hierarchy. Among the resting-state circuitry, the DMN represents a gold interface to describe the disease progression (Greicius et al., 2004; Koch et al., 2012), because the anatomical regions taking part in this network strictly resemble
the pattern of pathological build-up of peptidic biomarkers (Sperling et al., 2009), and it is also strongly correlated with cognitive deficit (Jones et al., 2011). The idea of AD as a “disconnection syndrome” is supported by both functional (Buckner et al., 2008; Greicius et al., 2004) and structural (Villain et al., 2008) data. Given this background, cognitive stimulation assumes a completely different meaning, because it is not about restoring cognition as a computational process of brain areas anymore, but it is about restoring cognition as a computational process of brain circuits. If cognitive deficits in AD are interpreted as a network disruption, it means that the neuropsychological difficulties emerge because of two distant areas not working synchronously any longer. Targeting only one hub of this circuit is not beneficial because it does not consider the circuit as a whole.

The DMN can be normally measured with an ICA or with a seed-based correlation of brain activity. None of these methods is probably adequate for test-retest comparisons: ICA permits the identification of the complete network with a focus on its entirety, whereas a seed-based approach lacks an explorative approach. A simple resting-state analysis of whole-brain BOLD signal can be implemented more conveniently, and, without violating the initial intentions to use a paradigm for both design of the training and measurement of the effects, test-retest differences in BOLD signal can be interpreted in relation to the DMN.

Normally any cognitive exercise would require the co-activation of multiple areas. For example, learning new face-name associations results in a mediotemporal and temporal activation, with additional components located in the parietal and frontal lobes and in basal ganglia (Sperling et al., 2001). Other tasks may evoke simpler responses, for example a standard Posner task only requires computations in the temporo-parietal junction and superior parietal lobe (Vossel et al., 2009), without projection to distant areas. The DMN starts showing the first signs of disconnection of the main hubs in the MCI stage of AD (Qi et al., 2010) and, as the disease progresses, a down-regulation of the posterior component of the DMN is observed, with a progressive decrease in activity of parietal, temporal and midline structures, paralleled by an up-regulation of the anterior component of the DMN, in ACC and prefrontal clusters (Jones et al., 2011). Provided that these increments and decrements are due to altered connectivity, a non-pharmacological restoration of these parameters could be achieved only if the cognitive exercises are specifically designed to target distant areas at the same time. This means that the tasks must tap synchronous multiple processing and must specifically target the hubs where a down-regulated synchronicity is observed. In comparison, any other cognitive stimulation can be considered unstructured, because either it was not designed with the purpose of restoring connectivity, or because it stimulates general connectivity without focusing on DMN decline. For example, even though it is not a scientific publication, the
instruction booklet of the Nintendo® DS Brain Training videogame features an introduction written by Dr Ryuta Kawashima, who also contributed to other papers quoted in this work (Kawashima et al., 2005; Nouchi et al., 2012; Takeuchi et al., 2010; Takeuchi, Taki, Hashizume, et al., 2011; Takeuchi, Taki, Sassa, et al., 2011; Uchida & Kawashima, 2008). Dr Kawashima reported that quick solving of simple arithmetic problems and reading aloud are two tasks that activate a big set of areas, and, according to him, the amount of simultaneous activation would be proportional to the long-term benefit of stimulation. Obviously in the manual he did not report the exact list of areas showing an effect, preferring to attach instead only some images depicting brain activation, not easily interpretable from the small pages. It is curious to note that Dr Kawashima, despite having nicely illustrated the effects of Brain Training tasks in the game booklet back in 2006, wrote the following sentence in one of his latest papers: “Yet in all honesty, the beneficial effects of the brain training games have little scientific basis” (Nouchi et al., 2012). One of the two tasks used as examples in the game manual was described by his team in a prior paper. Reading aloud was reported to activate the thalamus, occipital, temporal and motor areas bilaterally (Miura et al., 2003). Structural and functional connectivity displayed by these areas could be potentially enhanced by the Brain Training task involving reading aloud, but it would be not specific for AD related decline in regional connectivity. For this reason, different and specific exercises are needed. A recent study showed that when a stimulation programme is designed within a theoretical framework, and thus has a fair degree of specificity, it is more effective than an unstructured stimulation (Jelcic et al., 2012). However, specific stimulation has never been tested in terms of brain function.

Takeuchi et al. (2011) measured differences in functional connectivity from a seed-region (determined by the activation peak during a criterion task) before and after 5 days of speed-of-processing training. Increased connectivity was found between the seed (located in the left perisylvian region) and bilateral calcarine-lingual areas. The PCC was chosen as seed region by Voss and colleagues (2012) to compute functional connectivity in a group of young adults before and after 20 hours of training with the "shooter" videogame Space Fortress. Increased connectivity was found between PCC and left hippocampus (not significant at cluster level, though) and interpreted as a consolidation of part of the DMN. Resting-state whole-brain CBF was measured with perfusion MRI in a sample of young adults by Mozolic, Hayasaka & Laurienti (2010). After 8 weeks of training based on individually-shaped attentional exercises, increased rCBF was found in the right inferior frontal cortex compared to the placebo treatment. With a similar methodology, Mazoyer et al. (2009) reported resting-state rCBF increases in right ventromedial prefrontal and superior temporal cortices, right peristriatal
cortex and right pulvinar using PET. Their training consisted of 30 minutes of inhibition of a logical rule.

Given this evidence, we created a protocol of CA based on the principles of the DMN deficits exhibited in the early stage of AD. We wanted to test the applicability of this programme and its effectiveness in the modulation of brain function in the absence of pathology.

The DMN does not display any ceiling effects in later adulthood, because non-pathological, yet qualitatively similar alterations of the DMN are observed with healthy ageing. Decreased connectivity was reported in the posterior component with minor alterations in the anterior component, accompanied by a loss of inter-componential connectivity (Wu et al., 2011). Testing the effect of a CA programme on the DMN in healthy adults would permit further explanations of whether it triggered an augmented connectivity in the desired direction.

5.1.1.2 METHODS

5.1.1.2.1 PARTICIPANTS

Healthy adults (age≥50 y.o.) were enrolled from the archipelago of Venetian islands, and particularly from the island of Lido (Italy). Participants were split between experimental and control group and it was chosen to assign the complete first batch to the experimental group. Whenever a person was not available for participation, they were asked if they wanted to take part in the untreated control group. A second batch of participants were not specifically briefed about the procedure but were simply tested at baseline for assessment purposes and referred to a retest examination.

A full neurological examination and a comprehensive neuropsychological battery (described in Section 5.1.3.2.1) served as criteria to classify the person as suitable or not suitable. No specific medical exclusion criteria were set.

Twenty-nine healthy elderly adults were recruited in total. Scans were collected for 23 of these: sixteen were treated with the CA programme and 7 formed the control group. The approximate time spent for each session of training was 60 to 90 minutes. Characteristics of the samples are described in Table 5.1 and 5.2.
Ethical approval was obtained for Experiments 4 to 10 from the joint ethical committee of the research institute IRCCS San Camillo, Venice, Italy. Informed consent was obtained from all participants for the complete study. The local G. P. s in charge of the participants were informed about their patient taking part in this project.

---

Table 5.1 Demographic characteristics of the sample.

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Education</th>
<th>Gender (M/F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete Sample</td>
<td>67.6 (1.6)</td>
<td>10.3 (0.9)</td>
<td>10/13</td>
</tr>
<tr>
<td>Experimental Group</td>
<td>66.7 (1.7)</td>
<td>10.6 (1.2)</td>
<td>7/9</td>
</tr>
<tr>
<td>Control Group</td>
<td>69.7 (3.9)</td>
<td>9.6 (1.6)</td>
<td>3/4</td>
</tr>
</tbody>
</table>

Table 5.2 Cognitive characteristics of the samples at baseline.  

<table>
<thead>
<tr>
<th>Test</th>
<th>Experimental Group</th>
<th>Control Group</th>
<th>Cut-Off</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Corrected Score</td>
<td>N</td>
</tr>
<tr>
<td>MMSE †</td>
<td>22</td>
<td>28.59</td>
<td>7</td>
</tr>
<tr>
<td>Raven Matrices</td>
<td>22</td>
<td>31.90</td>
<td>7</td>
</tr>
<tr>
<td>Digit Cancellation</td>
<td>22</td>
<td>54.44</td>
<td>8</td>
</tr>
<tr>
<td>Stroop Test - Time</td>
<td>22</td>
<td>16.52</td>
<td>8</td>
</tr>
<tr>
<td>Stroop Test - Errors</td>
<td>22</td>
<td>0.50</td>
<td>8</td>
</tr>
<tr>
<td>Phonemic Fluency</td>
<td>22</td>
<td>36.20</td>
<td>8</td>
</tr>
<tr>
<td>Semantic Fluency</td>
<td>22</td>
<td>46.23</td>
<td>8</td>
</tr>
<tr>
<td>Token Test</td>
<td>21</td>
<td>35.07</td>
<td>7</td>
</tr>
<tr>
<td>Confrontational Naming †</td>
<td>22</td>
<td>19.68</td>
<td>7</td>
</tr>
<tr>
<td>Similarities ††</td>
<td>21</td>
<td>11.50</td>
<td>7</td>
</tr>
<tr>
<td>Digit Span Forward</td>
<td>22</td>
<td>5.59</td>
<td>8</td>
</tr>
<tr>
<td>Digit Span Backwards †</td>
<td>22</td>
<td>3.64</td>
<td>8</td>
</tr>
<tr>
<td>Paired Associates Learning</td>
<td>22</td>
<td>13.48</td>
<td>8</td>
</tr>
<tr>
<td>Short Story Recall</td>
<td>22</td>
<td>20.98</td>
<td>7</td>
</tr>
<tr>
<td>Visuospatial Span</td>
<td>22</td>
<td>5.09</td>
<td>8</td>
</tr>
<tr>
<td>Visuospatial Supraspan</td>
<td>21</td>
<td>20.90</td>
<td>8</td>
</tr>
<tr>
<td>Rey Figure - Copy</td>
<td>22</td>
<td>33.83</td>
<td>8</td>
</tr>
<tr>
<td>Rey Figure - Recall</td>
<td>22</td>
<td>18.69</td>
<td>8</td>
</tr>
</tbody>
</table>

* Performance under cut-off score
† Raw score is indicated
†† WAIS-R score is indicated

147 The cognitive scores refer to the full sample that took part in Experiment 6.


5.1.2.2 MATERIAL

5.1.2.2.1 PROGRAMME OF COGNITIVE STIMULATION

As only a few studies have been carried out with a full awareness of the effects of the programme on neurocognitive variables (Whitlock et al., 2012), the creation of these exercises was based on a model of AD-based decline in regional connectivity, with a concomitant involvement of distant areas, of crucial relevance in early-stage AD.

Exercises were grouped into 5 modalities according to their main feature: semantic retrieval, logical reasoning, proper names retrieval, speed of processing, plus an extra “mind-twister” exercise administered as end of each daily session and consisting of a single trial. Twenty sessions of exercises were created.

All tasks were administered through the E-Prime® Software, Version 2.0. Accuracy and RT were recorded for all trials. The first session was slightly shorter than the others, and from the second session on the material consisted of old and new trials mixed (1/3 old, and 2/3 new). The programme was not tailored on individual performances but was kept fixed. The repetition trials were selected at random before the beginning of the training. The only task in which stimuli were tailored on individual performance was the Proper Names Retrieval.

Verbal trials consisted of black-coloured words and/or sentences on a white background. No participant complained about a too-small font size. Similarly, image trials consisted of black-coloured images on a white background. All original images were characterised by a realistic style and were not cartoonish or caricature-style. The Snodgrass-Vanderwart Pictures Set (Snodgrass & Vanderwart, 1980) was used as source of stimuli, together with websites offering free educational material. Some images were specifically created for the purposes of the training.

We attempted to subjectively judge the level of difficulty of every single trial in order to obtain a programme of gradually increasing difficulty.

The participants were asked to respond to the trials by pressing a button on the keyboard. The keys necessary for the tasks were the spacebar and 1, 2, 4, 5 located on the keypad on the right hand-side. All stimuli appearing on the screen were accompanied by a number, and to respond to the trial the participants were asked to key in the number associated with the stimulus they wanted to choose.

Instructions were always presented before the beginning of all tasks, for all sessions. Apart from the Speed of Processing exercises, the stimuli remained on the screen until the
participants chose to respond. Although the training was centred on accuracy, rather than RT, the participants were asked to respond as soon as they had solved a trial.

All verbal stimuli were in Italian language and stimulation was carried out in Italian.

5.1.1.2.2.1.1 TASKS BASED ON SEMANTIC RETRIEVAL
Semantic knowledge has been described as represented by an extensive network diffused to many cortical areas of the parietal, temporal and frontal lobe (Binder, Desai, Graves, & Conant, 2009). We created exercises based on visual and verbal stimuli that could cover a big range of semantic content, with simultaneous computational requirements related to retrieval mechanisms, ambiguity and interference, working memory, and inhibition. The neural substrate normally related to these extra components is located in key areas of the DMN, namely the mediotemporal formation, the ACC and the prefrontal cortex.

- RULE THE ODD ONE OUT 1
Four words were presented on the screen. Three of them belonged to the same semantic category, and one was unrelated to the other three. This alien word had to be identified and selected.

- RULE THE ODD ONE OUT 2
Four images were presented on the screen. Similarly to the previous exercise, three of them could be grouped in one category, while the fourth one did not fit with this category and had to be selected.

- DOUBLE CATEGORY
Four words were presented on the screen as with the “Rule the Odd One out 1” exercise, but one of the words had a double meaning. One of the two meanings was associated with two of the other words, whereas the second meaning was associated with the fourth stimulus, for example, RUBBER, RULER, PENCIL, and KING. RULER has two meanings; its first meaning is associated with RUBBER and PENCIL, and its second meaning is associated with KING. This semantic interference must be resolved by understanding which one is the meaning to pick up in order to obtain a category with 3 members and allow the completion of the trial by the final selection of the odd one out (KING, in this case).
- **SEMANTIC INHIBITION**

A sentence was presented on the screen and the participant had to read it carefully and keep it in mind. When they felt confident they could remember it, they had to press a key and 4 images were presented; the prior sentence was consistent with 3 of these images but made no sense with 1 of them, which had to be selected. In this task participants were not required to choose according to a "positive" feature (a match between cue and stimulus), but rather according to a "negative" feature (the only stimulus that did not match with the cue). As this was the last task of a sequence of exercises in which "positive" features guide the response, it was necessary to inhibit the temptation to respond according to the detection of the cue-stimulus match.

5.1.1.2.2.1.2 **TASKS BASED ON LOGICAL REASONING**

Evidence coming from fMRI research has revealed that the three sequential phases of deductive reasoning are associated to three different neural contributions: after the initial examination of the material, a phase based on evidence-gathering and integration follows, with prefrontal and occipito-temporal involvement, followed in turn by the validation phase, in which a prefrontal and parietal circuit is recruited (Fangmeier, Knauff, Ruff, & Sloutsky, 2006). Based on this evidence, we created exercise of logical deductive reasoning in which a strong semantic processing was required to understand the correct answer.

- **SEQUENCE COMPLETION 1**

Two words were presented on the left side of the screen, connected by an arrow. A relation existed between these two words. The same relation was repeated for other two words in the middle of the screen, similarly connected by an arrow; however, one of these was missing (and substituted by a question mark). Two other words were located on the right side and one of them represented the correct answer to complete the sequence. Both choices were semantically related with the available word, but only one of them would ensure that the sequence is respected following the same relation between the two guide-words on the left.

- **SEQUENCE COMPLETION 2**

This task was as the "Sequence Completion 1" task, but this time no words were presented and the deductive reasoning had to be performed based on image stimuli only.

- **SENTENCE COMPLETION**

A sentence was presented in the top half of the screen, with one element missing (the subject, the verb, or a complement). Four possible choices were available in the bottom half of the
screen and only one represented the correct answer. The response had to be found by a process of elimination in which one or two options could be logically ruled out, and very often a semantic feature was crucial to make the final choice.

- **SCENE COMPLETION**
A complex scene was presented in the middle of the screen, in which an element had been removed (the main element, the background or a detail). Four smaller images were presented at the four corners of the screen and the participant had to select the image that best completed the scene. Similarly to the “Sentence Completion”, the final choice had to be made based on logical and semantic elements.

5.1.2.2.1.3 **TASKS BASED ON RT**
RT is considered a basic executive component contributing significantly to cognitive efficiency. Evidence has shown how faster participants performing a sort of go-no-go task display a larger parietal and a smaller frontal activation than slower adults (Rypma et al., 2006), in line with the idea that faster cognitive performance is mirrored by a more efficient neural computation, interestingly described as an anterior down-regulation and posterior up-regulation.
We created 4 simple tasks in which we specifically asked the participants to press a key as rapidly as possible. Stimulus onset asynchrony was variable to avoid facilitation of response.

- **RESPOND TO THE A**
A fixation cross was presented, followed by a capital A, appearing in the centre of the screen. The participant had to press a button as quickly as possible when they saw the stimulus.

- **RESPOND TO THE SQUARE**
This was similar to the “Respond to the A” task, but the stimulus to respond to was a blue square.

- **RESPOND TO A & B**
After the fixation cross a capital A, or B, could appear in the centre of the screen. The participant had to press one of two different keys as quickly as possible depending on which stimulus was presented.

- **RESPOND TO BLUE & RED SQUARES**
This task was similar to the previous one, but the differential response was driven by the colour of the square.
5.1.1.2.1.4 TASKS BASED ON PROPER NAMES RETRIEVAL

Recalling the name and/or surname of a well-known person is a difficulty that older people very often complain about, and it has been objectively described in MCI together with an impairment of other forms of proper name retrieval (Ahmed, Arnold, Thompson, Graham, & Hodges, 2008). A factor analysis revealed that phonological access problems seem to underlie the concurrent tip-of-the-tongue phenomenon with proper names in MCI individuals (Juncos-Rabadan, Rodriguez, Facal, Cuba, & Pereiro, 2011). Proper names for people or for topographical/geographical places are processed by a network of areas that has not been yet fully understood (Yasuda, Nakamura, & Beckman, 2000). We chose to include this task as a "symptomatic" exercise for participants, as it is a very common complaint manifested during neurological examination.

A brief written description (1 or 2 sentences) of a very famous person/place was presented on the screen. The participant had to read it and communicate the name corresponding to the description to the tester. If the participant could not retrieve the name, additional details about the person/place were given. If the participant still could not come up with the answer, this was provided by the tester, and the same trial would be asked on the following day in addition to the new trials. If the participant could still not recall the answer on the following day, the first letter of the name or first letter of name and surname were given as a cue. If the response was not given, additional letters were given, up to a total of four cues. If the participant still could not give the answer after 4 cumulative cues, this was provided by the tester. Whenever the response could not be given without the help of a cue, the trial was re-asked in the next session. Since this approach is errorless, no information about accuracy or RT was registered.

5.1.1.2.1.5 MIND-TWISTER FILLER

A single-trial task was created, based on lexical-semantic processing. This task was not part of the programme, but served as a filler task in case the tester needed to focus on a problematic trial with one participant and could not dedicate much time to other participants. It also served as a final task for participants that joined the experiment together (e.g. husband and wife) but one of them finished earlier than the others. It was preferentially left at the end of the session, as it was very stimulating and participants left the lab and the hospital in a positive mood. Sometimes the response was given immediately, sometimes more time was needed.

Three words were presented on the screen, completely unrelated with each other. The task was to find a fourth word that was verbally related to each of the stimuli in specific expressions and ways of saying, in which the global meaning was beyond the simple apposition
of the terms: for instance a potential trial in English could be: TWISTER, CHEEK, and MOTHER. The response is TONGUE, as TONGUE-TWISTER, TONGUE-IN-CHEEK and MOTHER TONGUE represent three semantic constructs independent from the literal combination of cue and target.

5.1.1.2.2.2 MRI SCANS ACQUISITION
Echo planar T2* weighted MRI images were acquired on a 1.5T Philips® Achieva system (TR=2s, TE=50ms, flip angle 90°, voxel dimensions 2.00x2.00x3.00mm³, field of view 230 mm). Two 120-volume runs of 20 contiguous axial slices were acquired in ascending order. Each run lasted approximately 4 minutes. Prior to proper scans acquisition 20s of dummy scans were acquired to allow the scanner to reach equilibrium.

During the scans, no specific instruction was given to the participants. They were only asked to remain as still as possible for the full duration of the scan and avoid any movements. A safety buzzer was given to them in case of required assistance. Similarly to Han et al. (2012), we chose not to give our participants any instructions during the resting-state scan because we assumed that a within-subject design would automatically remove any artefacts due to subjective variables having a significant impact on the scan (e.g. eyes open vs. eyes closed, mental activity performed during the scan, sleeping vs. remaining awake).

5.1.1.2.2.3 MRI SCANS PREPROCESSING AND ANALYSIS
Data were preprocessed and analysed using SPM-8 (Wellcome Trust Centre for Neuroimaging, London, UK).

Functional preprocessing included slice timing, realignment for motion correction, normalisation to stereotactic space and smoothing. The two runs were initially slice-timed in ascending order (slice 1 to 20) by using slice 10 as reference. Each of the two sessions of volumes was then realigned independently: a mean volume was created as reference, and 6 parameters of head motion were estimated in relation to the mean by using 4th Degree B-Spline interpolation methods. Images were then resliced based on these parameters. The volumes were then normalised using the first realigned volume of the first session as source image to match the default EPI template available in SPM-8, and setting the voxel dimensions to 2x2x3mm³. Finally, the images were spatially smoothed with a 6x6x9mm³ width at half maximum isotropic Gaussian kernel to fix possible inter-subject differences that had persisted after normalisation.
First level analysis aimed to create contrast images for each participant (post- vs. pre-treatment and pre- vs. post-treatment), covariating for movement vectors created during realignment, and applying a high-pass filter of 550ms. One-sample *t* tests were separately run for both sets of contrast images to assess changes in brain function in the sole experimental group. A two-sample *t* test was then run to compare the experimental and the control groups. For all analyses, a *p* value of 0.005 (uncorrected) was selected, and only clusters of at least 15 contiguous voxels were considered. This statistical threshold was previously judged as the best compromise in test-retest designs in samples of MCI patients (Fouquet et al., 2009). All reported clusters in a priori predicted brain areas were corrected for spherical small-volume (8mm diameter), and *p* values equal to or smaller than 0.05 were considered significant. Age and education were used as covariates.

Peak coordinates were transformed in Talairach stereotaxic space and results were interpreted using the Talairach Daemon client (www.talairach.org/client.htm). Peak values were processed using the “Single Point” database search option. Peak coordinates are reported in the results section.

**5.1.1.3 PROCEDURE**

The stimulation lab consisted of four separate workspaces equipped with desk and computer, placed at the four corners of the room and facing against the wall, to allow for individual work. Simultaneous treatment for multiple participants was possible. The tester was present in the room for the full time of the stimulation and no other people had access to the lab during this time. Participants were instructed to carry out their individual work in silence and were allowed to ask for assistance in case of difficulty or ambiguity. The tester supervised the training and in case of help needed, he did not provide specific help but rather a refresh of the instructions of the task with emphasis on the problematic trial. As each exercise consisted of a separate E-Prime® file, the tester was in charge of the start-up of all tasks for all participants, who were instructed to wait for the tester at the end of their performance in a task. The order of the tasks was randomly chosen by the tester, to avoid the scenario of multiple participants finishing a task at the same time and waiting for too long before starting a new one, and to save sufficient time to allow all participants to perform the Retrieval of Proper Names Task with discretion (and in a low voice) together with the tester. Participants were also allowed to take advantage of the numerous pauses between and within tasks.

As the time of training was scheduled on a day-by-day basis by participant and tester together, participants were allowed to arrive to the lab at a flexible (morning to evening), and variable
(e.g. Monday at 8am and Tuesday at 2pm) time, even when another participant had already started their session. Out-of-phase sessions were possible, with simultaneous training of participants who were at a different stages (e.g. participant 1 performing session 4/20 together with participant 2 performing session 7/20).

5.1.1.4 RESULTS

Baseline scans were collected between 1 and 63 days (mean=17.8 days; SEM=4.3) prior to the beginning of the training. Retest scans were instead constantly collected immediately after the end of the programme (max 2 days). Based on this, the retest scans were collected on average 41.9 days after the baseline (SEM=4.5; range=20-86). In the control group this interval was 52.6 days (SEM=9.4; range 27-107). This value did not differ between the 2 groups ($t_{21}$=-1.16; $p>0.05$). The moment of the scan was kept as constant as possible within each participant, as for both weekday and time.
Parameters of translation were visually checked. As ours was a within-sample design, we wanted to make sure that any movement inside the scanner could not have influenced the analysis by being larger in one of the two sessions. We therefore selected for further analyses those participants who displayed a 1.5mm movement peak in one of the three axes for at least three times during the two echoplanar scans of each session. No problematic translation pattern was found.

Parameters of rotation were not analysed because rotation often displays a cumulative trend which does not allow for a linear analysis.

The one-sample t test revealed an increase in BOLD signal in bilateral areas of the temporal, parietal and frontal lobe, with additional significant clusters in the cerebellum and thalamus (see Table 5.3 and Figure 5.2). The biggest increase was found in posterior portions of the brain, areas that are part of the posterior component of the DMN, such as cortical clusters of the superior and middle temporal gyri bilaterally, the precuneus bilaterally, and mediotemporal-limbic structures (left parahippocampal gyrus and right amygdala). One single increase was found in prefrontal regions. No significant decreases were found from baseline to retest.

Fig. 5.2 Increased activation in the healthy subgroup following treatment. In the three slices it is possible to see the involvement of the precuneus, right temporal-limbic areas, and subcortical clusters of thalamus and external capsule.
The two-sample t test revealed a well-defined pattern of increased activation in the posterior half of the brain for the experimental group. Nine clusters were found to be significant in important DMN areas: the right PCC, the parahippocampal gyrus bilaterally, a subcortical limbic cluster adjacent to the left hippocampus (yielding the highest Z score), and a WM cluster adjacent to the inferior portion of the left parahippocampal gyrus. Two clusters were found in the occipital lobe and one cluster emerged in the WM neighbouring the superior portion of the left insula (Table 5.4; Figure 5.3). No cluster was found in the frontal lobe, not even by increasing the p value to a more liberal level (Figure 5.4).

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/ Right</th>
<th>N° voxels</th>
<th>Z value at Local Maximum</th>
<th>Tailarach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-lobar Extra-Nuclear</td>
<td>L WM</td>
<td>121</td>
<td>4.23</td>
<td>-16 -5 18</td>
</tr>
<tr>
<td>(Internal Capsule)</td>
<td></td>
<td></td>
<td>3.71</td>
<td>-14 -4 7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.22</td>
<td>-8 -6 1</td>
</tr>
<tr>
<td>Middle Temporal Gyrus</td>
<td>R WM</td>
<td>102</td>
<td>4.14</td>
<td>55 -22 -8</td>
</tr>
<tr>
<td>Temporal Sub-Gyral</td>
<td>R WM</td>
<td>3.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle Temporal Gyrus</td>
<td>R WM</td>
<td>3.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Precuneus</td>
<td>R 7 GM</td>
<td>31</td>
<td>4.10</td>
<td>4 -48 59</td>
</tr>
<tr>
<td>Thalamus</td>
<td>L GM</td>
<td>135</td>
<td>3.98</td>
<td>-14 -28 -10</td>
</tr>
<tr>
<td>Pulvinar</td>
<td>L GM</td>
<td>3.56</td>
<td>-18 -28 13</td>
<td></td>
</tr>
<tr>
<td>Ventral Posterior Lateral Nucleus</td>
<td>L GM</td>
<td>3.28</td>
<td>-18 -23 2</td>
<td></td>
</tr>
<tr>
<td>Precuneus</td>
<td>R WM</td>
<td>3.91</td>
<td>2 -64 37</td>
<td></td>
</tr>
<tr>
<td>Precuneus-Cuneus</td>
<td>L WM</td>
<td>3.85</td>
<td>-4 -71 27</td>
<td></td>
</tr>
<tr>
<td>Superior Temporal Gyrus</td>
<td>R WM</td>
<td>103</td>
<td>3.79</td>
<td>34 3 -22</td>
</tr>
<tr>
<td>Amygdala</td>
<td>R GM</td>
<td>3.15</td>
<td>24 -5 -22</td>
<td></td>
</tr>
<tr>
<td>Precentral Gyrus</td>
<td>R 6 GM</td>
<td>25</td>
<td>3.61</td>
<td>53 -2 40</td>
</tr>
<tr>
<td>Superior Temporal Gyrus</td>
<td>R 38 GM</td>
<td>27</td>
<td>3.59</td>
<td>57 9 -12</td>
</tr>
<tr>
<td>Superior Temporal Gyrus</td>
<td>L WM</td>
<td>60</td>
<td>3.35</td>
<td>-32 -22 -20</td>
</tr>
<tr>
<td>Middle Frontal Gyrus</td>
<td>L WM</td>
<td>17</td>
<td>3.42</td>
<td>-36 8 39</td>
</tr>
<tr>
<td>Cuneus</td>
<td>L 23 GM</td>
<td>17</td>
<td>3.30</td>
<td>-2 -75 10</td>
</tr>
<tr>
<td>Parahippocampal Gyrus</td>
<td>L WM</td>
<td>19</td>
<td>3.27</td>
<td>-20 -12 -14</td>
</tr>
<tr>
<td>Superior Temporal Gyrus</td>
<td>L WM</td>
<td>21</td>
<td>3.23</td>
<td>-48 8 1</td>
</tr>
<tr>
<td>Cerebellum - Declive of Vermis</td>
<td>L GM</td>
<td>17</td>
<td>3.20</td>
<td>-2 -65 -19</td>
</tr>
<tr>
<td>Cingulate Gyrus</td>
<td>L GM</td>
<td>27</td>
<td>3.20</td>
<td>-6 10 36</td>
</tr>
<tr>
<td>Temporal Sub-Gyral</td>
<td>L WM</td>
<td>27</td>
<td>3.16</td>
<td>-34 -67 13</td>
</tr>
<tr>
<td>Temporal Sub-Gyral</td>
<td>L WM</td>
<td>3.06</td>
<td>-28 -71 24</td>
<td></td>
</tr>
<tr>
<td>Middle Temporal Gyrus</td>
<td>L WM</td>
<td>21</td>
<td>2.92</td>
<td>-50 -39 -2</td>
</tr>
</tbody>
</table>
Table 5.4 Areas displaying increased activation in the experimental group compared to the controls.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/Right</th>
<th>BA</th>
<th>No. of voxels</th>
<th>Z value at Local Maximum</th>
<th>Talairach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limbic Sub-Gyral</td>
<td>L</td>
<td>WIM</td>
<td>71</td>
<td>4.02</td>
<td>-19 -29 -3</td>
</tr>
<tr>
<td>Sublobar Extra-Nuclear</td>
<td>R</td>
<td>WIM</td>
<td>130</td>
<td>3.35</td>
<td>34 -5 7</td>
</tr>
<tr>
<td>Putamen</td>
<td>R</td>
<td>GM</td>
<td>3.31</td>
<td>23 -12 -2</td>
<td></td>
</tr>
<tr>
<td>Posterior Cingulate</td>
<td>R</td>
<td>WIM</td>
<td>127</td>
<td>3.73</td>
<td>15 -62 10</td>
</tr>
<tr>
<td>Lingual Gyrus</td>
<td>R</td>
<td>WIM</td>
<td>2.95</td>
<td>20 -66 -1</td>
<td></td>
</tr>
<tr>
<td>Parahippocampal Gyrus</td>
<td>R</td>
<td>WIM</td>
<td>33</td>
<td>3.37</td>
<td>13 -37 3</td>
</tr>
<tr>
<td>Parahippocampal Gyrus</td>
<td>L</td>
<td>GM</td>
<td>28</td>
<td>3.30</td>
<td>-28 -32 -15</td>
</tr>
<tr>
<td>Sub-lobar Extra-Nuclear</td>
<td>L</td>
<td>WIM</td>
<td>18</td>
<td>3.27</td>
<td>-32 -3 12</td>
</tr>
<tr>
<td>Temporal Sub-Gyral</td>
<td>L</td>
<td>WIM</td>
<td>23</td>
<td>3.15</td>
<td>-48 -37 -10</td>
</tr>
<tr>
<td>Cuneus</td>
<td>L</td>
<td>WIM</td>
<td>16</td>
<td>3.05</td>
<td>-24 -7 7 10</td>
</tr>
<tr>
<td>Lingual Gyrus</td>
<td>L</td>
<td>WIM</td>
<td>48</td>
<td>2.95</td>
<td>-18 -57 -4</td>
</tr>
<tr>
<td>Lingual Gyrus</td>
<td>L</td>
<td>GM</td>
<td>16</td>
<td>2.92</td>
<td>-16 -64 -1</td>
</tr>
</tbody>
</table>

Fig. 5.3. Two-sample t test comparing the two subgroups of healthy participants. The sagittal view shows the contribution of cuneus and PCC in its retrosplenial portion, whereas the posterior parahippocampal gyrus is illustrated in the coronal section. The axial slice shows a cluster that extends from the external to the internal capsule, including the putamen.

5.1.1.5 DISCUSSION
The choice of a resting-state paradigm justifies the exploratory intent of our study and offers an interpretational model that accounts for the pathological decline in brain function observed at various clinical stages of AD. As a network-based organisation of brain function is hierarchically superior to an area-based organisation, it is believed to represent a more accurate theoretical framework. In a network-based structure, improvements in brain function reflect changes in overall and regional connectivity between distant hubs. Cheng et al. (2012) refer to this by using the Hebb's paradigm “neurons that fire together, wire together” not
applied to single cells connections, but to a larger scale of multiple-neurons network. Within this interpretational framework, analysis of functional connectivity would represent an ideal methodology to study the DMN. However, it is also possible to maintain a high-hierarchy interpretational model and use instead low-hierarchy investigation methods, as long as the interpretation of the findings is compatibly drawn from the initial high-hierarchy model (as done by Brown et al (2011)). We designed our training with the specific goal to tap multidimensional cognitive processing, and for this reason we assumed that our CA programme was specific for targeting multiple brain areas at the same time. Repeated synchronous activity of two or more hubs is the behavioural rationale for inducing neuroplasticity-based increases in connectivity and for this reason we expected test-retest differences due to training. The one-sample t test revealed that the experimental group displayed a modulation of resting-state BOLD signal that is consistent with the idea of DMN remodelling. In a healthy subgroup there is no pathological impairment of the DMN, and after our treatment a consolidation of areas belonging to the posterior DMN was seen. This indicates that it is possible to trigger neuroplastic mechanisms at older age and reverse the non-pathological decline in posterior DMN due to ageing. The two-sample t test investigated the experimental question by adding a control group: the comparison of the two healthy subgroups revealed how our CA programme increased activation in the posterior half of the brain. We can interpret this idea as further evidence that our programme modulated the posterior DMN subpart, and that appropriate CA can be protective against the onset of a pathology that disrupts brain function.

Ageing has been described to cause a reallocation of resting-state brain function of the anterior DMN (Jones et al., 2011), but in our study only limited benefits were observed in anterior hubs of the DMN. This indicates that our CA exercises were specific and modulated brain function with selectivity. We actually designed the programme to stimulate the DMN in its entirety, but only the posterior DMN seemed to respond. It is possible that the general benefit consisted of a better connectivity between the two sub-networks (Koch et al., 2012), rather than enhancements internal to one of the two or both the sub-networks. Unfortunately our methodology did not allow us to clarify this issue.
Interpreting our results in terms of increased functional connectivity means assuming an improvement in a statistical parameter associated to brain function. Even though structural and functional connectivity are related but independent (Damoiseaux & Greicius, 2009), an increase in functional connectivity would imply some sort of structural implementation. For this reason we studied changes in GM and WM volume due to CA.
5.1.2 EXPERIMENT 5 - EFFECTS ON BRAIN STRUCTURE

5.1.2.1 INTRODUCTION

The main corpus of reports focusing on training-dependent structural changes inside the brain comes from studies that have investigated the impact of perceptual-motor exercises through juggling training. Healthy young (Draganski et al., 2004) and old (Boyke, Driemeyer, Gaser, Buechel, & May, 2008) adults displayed a longitudinal GM change in the middle temporal part of the V5 area, plus an increase in the posterior section of left intraparietal sulcus (in the young group), in the nucleus accumbens and in the left hippocampus (in the older group) after 3 months of juggling exercises. This pattern of results has not always been replicated (Thomas & Baker, 2012), but nonetheless this is strong evidence that an old brain retains capacity for training-induced structural changes. Unfortunately, to our knowledge, no study has investigated the effects of a specific cognitive training on structural changes intended as induced neuroplasticity.

Engvig et al. (2010) designed a training programme based on the Method of the Loci and found that treated adults displayed a thicker right insular cortex compared to the control group, and this cortical thickening correlated with post-training performance in verbal recognition (but not with the improvement score). As already explained, teaching a strategy is not the best example of induced plasticity, but is a good starting point to demonstrate that it is possible to trigger structural changes even in old age. This area of research is flourishing and holds surprising results. May et al. (2007) tested for the first time an extremely brief paradigm of intense stimulation (they used repeated Transcranial Magnetic Stimulation), lasting five days only. The positive results they found induced other teams to use the same training regime applied to CA. A Japanese team has done it in samples of young adults. A 4-hour a day regime of working memory training was controlled with a five-day placebo stimulation: GM decreases were detected in the DLPFC and parietal cortex bilaterally, and in the left superior temporal gyrus. This decrease correlated with improvement in various cognitive aspects (Takeuchi, Taki, Sassa, et al., 2011). Five days of intense speed of processing training resulted in a GM decrease in the left superior temporal gyrus and in occipital/occipito-temporal areas bilaterally, and an increase in the right precentral gyrus (Takeuchi, Taki, Hashizume, et al., 2011). Longer training centred on visuospatial and n-back-like working memory was tested with DTI. After 2 months of daily exercises, students displayed increased voxel-based anisotropy in a cluster that connects the inferior parietal lobule to frontal structures and in another cluster shared between the frontal and parietal lobes (Takeuchi et al., 2010).
Sagi et colleagues (2012) instructed their participants to play the racing videogame Need for Speed trying to learn a track in order to obtain faster times. Compared to a control group in which the track changed every time (giving therefore no opportunity for learning), the experimental group showed diffusivity changes in various areas after just 2 hours of intervention. Young adults revealed reduced MD and increased FA in mediotemporal and limbic areas, suggesting that the aspects of plasticity might be not as “sluggish” as concluded by Lövdén et al. (2010). This would be particularly important also because long-lasting CA stimulations are time-demanding, and shorter paradigms could increase the amount of experimental evidence.

Based on this evidence, we measured changes in brain structure induced by our programme of CA. As neuroplasticity takes place only when changes of various nature in neural structure have taken place (Lövdén et al., 2010), finding changes in brain structure would provide supporting evidence to the idea that functional changes found in Experiment 4 were actually promoted by neuroplastic mechanisms. In the literature programmes of CA as remodelling of structural variables have only been studied in samples of young adults. This is the first study to look for whole-brain effects of specific CA on brain structure in late adulthood.

5.1.2.2 METHODS
The same participants described in Section 5.1.2.1 took part to this experiment. The characteristics of the sample are described in detail in Tables 5.1 and 5.2.
They were administered the material described in Section 5.1.2.2.1 as part of the same data collection following the previously-described methodology. Structural brain images were co-acquired together with the EPI scans. Machines and software were the same as in Experiment 4.

5.1.2.2.1 MRI SCANS ACQUISITION
Structural scans were recorded on the same machine as with the functional scans. T1-weighted sagittal MRI images were acquired in 3 dimensions with a Turbo Field Echo Sequence. Voxel dimensions were set at 1.1x1.1x0.6mm³, field of view was 250mm with a matrix of size 256 X 256 X 124. Complete acquisition time clocked at 4min and 27s, with a TR=7.4ms, TE=3.4ms and flip angle 8°. T2-weighted axial- and Fluid Attenuated Inversion Recovery (FLAIR) coronal scans were also acquired to evaluate vascular load and the presence of other possible abnormalities. T1-, T2-, and FLAIR-weighted images were acquired immediately after the echoplanar scans, in a unified series, about 25-minute long. The order of
the scans was kept constant, with sporadic changes due to sudden malfunctioning of the machine, which however did not prevent data collection.

5.1.2.2 MRI SCANS PREPROCESSING AND ANALYSIS

VBM preprocessing and analyses was carried out using SPM-8. Three-dimensional T1-weighted scans were segmented to separate GM and WM from cerebrospinal fluid according to a probabilistic tissue classification. During this step, GM and WM were normalised and modulated. Resulting maps were smoothed with a full width half maximum kernel of 6mm. Images were then compared within each group with separated paired-sample t tests, and between-groups change was analysed with a mixed design ANOVA. A p value of 0.005 uncorrected was chosen for the analyses of the structural scans as well, together with a cluster size of 15 contiguous voxels. A relative threshold of 0.8 was used to set a spatial constraint and ensure the analysis comprised only cerebral tissue in relation to the template covering the whole 3D space.

Structural GM coordinates were interpreted using the “Nearest Gray Matter” search of the Talairach Daemon. WM was instead interpreted by using the “Cube Range” option and progressively increasing the distance from the peak, from 0 to 5mm$^3$ until white matter was found (in case of more areas in the same volume, the structure with more voxels was chosen). Distance in millimeters was extracted and reported for both GM and WM. Findings indicating cerebellar WM and pontine clusters were discarded.

In addition, measures of total GM, WM and intracranial volume (TIV) were extracted at baseline and retest for all those who underwent MRI scanning, and were analysed inside the experimental group (paired t test) and comparing the experimental and the control group (mixed-design ANOVA).

A measure of TIV was chosen as covariate for the analyses of regional changes of brain structure in the mixed ANOVA. An average of the TIV at the two measured timepoints was computed and used.

5.1.2.3 PROCEDURE

The procedure was the same as in Experiment 4 (Section 5.1.1.3).

5.1.2.4 RESULTS

Neither the Kolmogorov-Smirnov nor the Shapiro-Wilk tests revealed deviations from normality in the distribution of the TIV, total GM and total WM scores, therefore parametric tests were used. Differences between baseline and retest values were found in none of the
global parameters (adjusting by age and level of education). Controls did not differ from the experimental subgroup for global levels of GM ($t_{21}=1.22; p>0.05$) or WM ($t_{21}=1.20; p>0.05$).

The one-sample $t$ tests revealed a set of GM changes: an increase was found in three frontal regions of the right hemisphere, in a large cluster comprising the left parahippocampal and fusiform gyri, and in four cerebellar clusters, three of which on the left side, showing peaks in five distinct areas (see Table 5.5 and Figure 5.5). A single decrease was found in a cluster located between cuneus and precuneus in the left hemisphere (see Table 5.6 and Figure 5.6).

Table 5.5 GM increases in the experimental group.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/ Right</th>
<th>BA</th>
<th>Voxels</th>
<th>Distance (mm)</th>
<th>Z value at Local Maximum</th>
<th>Talairach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebellum - Inferior Semi-Lunar Lobule</td>
<td>R</td>
<td>116</td>
<td>0</td>
<td>4.47</td>
<td>26</td>
<td>-81 -35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>3.19</td>
<td>18</td>
<td>-80 -36</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>3.15</td>
<td>14</td>
<td>-66 -45</td>
</tr>
<tr>
<td>Cerebellum - Pyramis</td>
<td>L</td>
<td>97</td>
<td>0</td>
<td>4.47</td>
<td>-26</td>
<td>-68 -30</td>
</tr>
<tr>
<td>Medial Frontal Gyrus</td>
<td>R</td>
<td>10</td>
<td>26</td>
<td>3.74</td>
<td>10</td>
<td>43 11</td>
</tr>
<tr>
<td>Cerebellum - Inferior Semi-Lunar Lobule</td>
<td>L</td>
<td>130</td>
<td>4</td>
<td>3.54</td>
<td>-34</td>
<td>-64 -46</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>3.25</td>
<td>-28</td>
<td>-76 -42</td>
</tr>
<tr>
<td>Cerebellum - Cerebellar Tonsil</td>
<td>L</td>
<td>4</td>
<td>3.23</td>
<td>-32</td>
<td>-49</td>
<td>-46</td>
</tr>
<tr>
<td>Cerebellum - Pyramis of Vermis</td>
<td>L</td>
<td>120</td>
<td>0</td>
<td>3.50</td>
<td>0</td>
<td>-69 -27</td>
</tr>
<tr>
<td>Cerebellum - Uvula</td>
<td>L</td>
<td>0</td>
<td>3.47</td>
<td>-8</td>
<td>-79</td>
<td>-31</td>
</tr>
<tr>
<td>Parahippocampal Gyrus</td>
<td>L</td>
<td>36</td>
<td>114</td>
<td>3.25</td>
<td>-26</td>
<td>-36 -13</td>
</tr>
<tr>
<td>Fusiform Gyrus</td>
<td>L</td>
<td>37</td>
<td>0</td>
<td>2.98</td>
<td>-36</td>
<td>-45 -13</td>
</tr>
<tr>
<td>Fusiform Gyrus</td>
<td>L</td>
<td>37</td>
<td>2</td>
<td>2.92</td>
<td>-44</td>
<td>-40 -13</td>
</tr>
<tr>
<td>Inferior Frontal Gyrus</td>
<td>R</td>
<td>47</td>
<td>29</td>
<td>3.17</td>
<td>20</td>
<td>20 -20</td>
</tr>
<tr>
<td>Orbital Gyrus</td>
<td>R</td>
<td>47</td>
<td>1</td>
<td>2.74</td>
<td>18</td>
<td>28 -22</td>
</tr>
</tbody>
</table>

Fig. 5.5 GM increase in the healthy subgroup: the cerebellar changes are illustrated, together with the parahippocampal gyrus and two frontal clusters.
Table 5.6 GM decreases in the experimental group.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/Right</th>
<th>BA</th>
<th>Nb voxels</th>
<th>Distance (mm)</th>
<th>Z value at Local Maximum</th>
<th>Talairach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precuneus-Cuneus</td>
<td>L</td>
<td>31</td>
<td>24</td>
<td>2</td>
<td>3.82</td>
<td>-12 -67 20</td>
</tr>
</tbody>
</table>

Fig. 5.6 GM decrease in the healthy subgroup: only one cluster was significant, located at the crossroads among dorsal portion of PCC, precuneus and cuneus.

Table 5.7 WM Increases in the experimental group.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/Right</th>
<th>BA</th>
<th>Nb voxels</th>
<th>Distance (mm)</th>
<th>Z value at Local Maximum</th>
<th>Talairach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-lobar Extra-Nuclear (Internal Capsule)</td>
<td>R</td>
<td>108</td>
<td>1</td>
<td>4.14</td>
<td>18 14 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.52</td>
<td>14 4 -5</td>
</tr>
</tbody>
</table>

Fig. 5.7 WM increase in the healthy subgroup.

One single WM change was detected, a single increase in the region located between the right caudate and the putamen (internal capsule) with two significant peaks (Table 5.7; Figure 5.7). No WM loss was reported.
The mixed-design ANOVA revealed a significant interaction in large set of areas. The experimental group showed larger GM increase compared to the control group in bilateral frontal and cerebellar areas, the left fusiform gyrus and right limbic structures (Table 5.8; Figure 5.8). Two additional clusters in the right parahippocampal gyrus and left hippocampus were found and survived small volume correction for a 7mm$^3$-sphere (1mm$^3$ less than the chosen threshold). For this reason we considered this increase as marginally significant. The inverted contrast revealed a differential GM decrease in the experimental group in the right precuneus and left putamen (Table 5.9; Figure 5.9). Four increases in WM were found after training, in the right occipital lobe and in three left subcortical clusters adjacent to the pulvinar and the globus pallidus (Table 5.10; Figure 5.10). Nothing emerged from the inverted contrast.

### Table 5.8 Training-induced GM increases in comparison to the control group.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/ Right</th>
<th>BA</th>
<th>No. voxels</th>
<th>Distance (mm)</th>
<th>Z value at Local Maximum</th>
<th>Tailarach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medial Frontal Gyrus</td>
<td>R 25</td>
<td>156</td>
<td>1</td>
<td>4.75</td>
<td>10 17</td>
<td>-16</td>
</tr>
<tr>
<td>Anterior Cingulate</td>
<td>R 24</td>
<td>0</td>
<td>3.76</td>
<td>2 25</td>
<td>-6</td>
<td></td>
</tr>
<tr>
<td>Anterior Cingulate</td>
<td>L 25</td>
<td>0</td>
<td>3.35</td>
<td>-2 13</td>
<td>-4</td>
<td></td>
</tr>
<tr>
<td>Parahippocampal Gyrus</td>
<td>R 35</td>
<td>99</td>
<td>2</td>
<td>4.04</td>
<td>24 -9 -23</td>
<td></td>
</tr>
<tr>
<td>Cerebellum - Inferior Semi-Lunar Lobule</td>
<td>R 193</td>
<td>0</td>
<td>3.94</td>
<td>28 -76 -40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rectal Gyrus</td>
<td>L 11</td>
<td>42</td>
<td>1</td>
<td>3.91</td>
<td>-8 15 -18</td>
<td></td>
</tr>
<tr>
<td>Fusiform Gyrus</td>
<td>L 37</td>
<td>207</td>
<td>1</td>
<td>3.89</td>
<td>-28 -38 -15</td>
<td></td>
</tr>
<tr>
<td>Cerebellum - Culmen</td>
<td>L</td>
<td>0</td>
<td>3.06</td>
<td>-24 -44 -18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior Frontal Gyrus</td>
<td>L 11</td>
<td>47</td>
<td>1</td>
<td>3.75</td>
<td>-34 48 -14</td>
<td></td>
</tr>
<tr>
<td>Fusiform Gyrus</td>
<td>L 37</td>
<td>37</td>
<td>0</td>
<td>3.59</td>
<td>-38 -57 -11</td>
<td></td>
</tr>
<tr>
<td>Cerebellum - Inferior Semi-Lunar Lobule</td>
<td>L 53</td>
<td>3</td>
<td>3.38</td>
<td>-20 -68 -45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebellum - Tuber</td>
<td>R 21</td>
<td>0</td>
<td>3.23</td>
<td>46 -71 -25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncus</td>
<td>R 20</td>
<td>29</td>
<td>2</td>
<td>3.12</td>
<td>26 -6 -37</td>
<td></td>
</tr>
</tbody>
</table>

### Table 5.9 Areas in which the decrease in GM was selectively steeper for the experimental group.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/ Right</th>
<th>BA</th>
<th>No. voxels</th>
<th>Distance (mm)</th>
<th>Z value at Local Maximum</th>
<th>Tailarach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Putamen</td>
<td>L</td>
<td>117</td>
<td>0</td>
<td>4.03</td>
<td>-22 0 -3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>3.59</td>
<td>-28 -6 -1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cuneus - Precuneus</td>
<td>R 31</td>
<td>97</td>
<td>2</td>
<td>3.68</td>
<td>6 -69 27</td>
<td></td>
</tr>
<tr>
<td>Precuneus</td>
<td>R 7</td>
<td>30</td>
<td>0</td>
<td>3.52</td>
<td>6 -58 38</td>
<td></td>
</tr>
</tbody>
</table>
Fig. 5.8 Group-by-timepoint interaction showing the areas where the change from baseline to retest was differentially increased among the treated participants: the right sagittal slice displays the involvement of the uncus and the parahippocampal gyrus, plus a large cerebellar cluster. The horizontal slice shows a bilateral increase in orbitofrontal GM and a large cluster on the left extending through the fusiform gyrus to the cerebellum. In addition it is possible to see a small cluster, replicated from a coronal perspective, corresponding to the hippocampus, which did not survive small volume correction.

Fig. 5.9 Representation of the areas showing larger GM decrease in the experimental group.

Table 5.10 WM increase in the experimental group in comparison to controls.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/Right</th>
<th>N- voxels</th>
<th>Distance (mm)</th>
<th>Z value at Local Maximum</th>
<th>Tailarach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-lobar Extra-Nuclear (Perithalamic)</td>
<td>L</td>
<td>74</td>
<td>3</td>
<td>4.65</td>
<td>-18 -27 11</td>
</tr>
<tr>
<td>Sub-lobar Extra-Nuclear (Posterior limb of Internal Capsule)</td>
<td>L</td>
<td>349</td>
<td>2</td>
<td>4.24</td>
<td>-20 2 -5</td>
</tr>
<tr>
<td>Sub-lobar Extra-Nuclear (Anterior limb of Internal Capsule)</td>
<td>L</td>
<td>68</td>
<td>0</td>
<td>3.36</td>
<td>-10 8 -2</td>
</tr>
<tr>
<td>Lingual Gyrus</td>
<td>R</td>
<td>31</td>
<td>0</td>
<td>3.17</td>
<td>22 -60 0</td>
</tr>
</tbody>
</table>
5.1.2.5 DISCUSSION

Our programme induced GM and WM changes in a sample of healthy elderly adults. This was visible from the results obtained with both one-sample t tests and mixed-design ANOVAs, and the results obtained with the two statistical techniques displayed a high degree of match, revealing a training-induced GM increase in the left fusiform gyrus, and in frontal, limbic and cerebellar lobes bilaterally. A GM decrease was observed in the precuneus, while WM increase emerged in large left subcortical clusters.

Structural changes were expected because significant increases in brain function are the result of changes in the neural substrate. Within-between mixed-design DTI has proved to be extremely sensitive to modification emerged over just 2 hours (Sagi et al., 2012), and for this reason WM selective changes were expected. We did not find any trend of WM increase, but a pattern of positive GM corroboration. It is arguable that people with no on-going neurodegenerative process have no major baseline WM alterations. For this reason engagement in specific DMN-based CA would not result in any specific WM increase, but rather in a qualitatively different event. It is possible to define cortical GM as the extremes of long-distance connections. Based on this the findings can be interpreted as a healthy system that gets even healthier with our CA programme. Although no substantial WM changes are displayed, complimentary evidence of GM changes can be interpreted as a function of our hypothesis and can support it.

Increase/decrease in GM/WM is the result of one or more biological mechanisms that should be identified or at least speculated. Changes in GM may be accounted for by other phenomena than just cellular/synapses gain/loss: for example the main cause of basal ganglia atrophy in AD is attributable to neuronal shrinkage (Rinne, Paljarvi, & Rinne, 1987). Based on this, if by
remote chance CA had increased perykarial size, a GM increase would represent a corroborative positive effect, speculatively observable only in the presence of intact WM. Obviously, as many cellular elements contribute to GM, any mechanism (or set of mechanisms) may be the underlying cause of GM increase, and the real aetiology of the change remains unknown.
5.1.3 EXPERIMENT 6 - EFFECTS ON COGNITIVE FUNCTIONS

5.1.3.1 INTRODUCTION

If a programme of CA triggers significant changes in brain structure and function, improvement in behavioural variables should be objectively visible or at least subjectively felt by the individual. Measuring this improvement is a challenge: validated cognitive tests are effective diagnostic tools but are weak in the experimental design (see the introduction of this chapter). Nonetheless it is interesting to overview the literature and fathom what changes in cognitive scores it is conceivable to aim for.

The Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) study (Jobe et al., 2001) is a large randomised trial that aimed to test three protocols of cognitive stimulation: memory, reasoning and speed of processing. The memory and the reasoning programmes were based on learning and application of strategies. As already described, strategy-based exercises are not the best example of plasticity-stimulating training as they are mainly structured on knowledge, not processes. This does not mean that the interventions have been ineffective: in fact, a positive effect was found for all experimental conditions after 8 or more training sessions, but improvements were limited to the trained domain (Ball et al., 2002).

Interestingly, when measuring other variables related to everyday life like medical expenditure and depressive symptoms, at least four papers found encouraging findings for speed-of-processing training only (for example Wolinsky et al. (2009)), suggesting that knowledge-based treatments may be of limited effect, and that the positive effect on cognition is due to practice and near transfer only. Speed of processing is often considered as a basic executive aspect of elaboration that can improve different types of cognitive processes. Edwards and his team (2005) administered a programme based on processing speed to a sample of healthy old; ten hours of stimulation led to improvement only in IADL, with no change in cognition.

When principles of cognitive plasticity are implemented in a programme of targeted cognitive stimulation, it is possible to extrapolate information about efficacy that is independent from intervenient factors and variables intrinsic to the sample. The Improvement in Memory with Plasticity-based Adaptive Cognitive Training (IMPACT) study is a double-blind trial specifically designed to assess the effect of exercises included in commercialised software (Brain Fitness, developed by Posit Science Corporation). Eight weeks of exercise of increasing challenge (one hour, 5 days every week) were administered to a large group of healthy elders. At retest the experimental group obtained a significant improvement in various neuropsychological tests compared to a control group provided with educational material: a positive effect was seen on speed of processing, verbal memory, and digit-span backwards (Smith et al., 2009). Some of
the benefits were lost at follow-up, suggesting that the efficacy of such a treatment is weak when a continuous reinforcement for the elderly is absent (Zelinski et al., 2011). The main assumption on which the training programme was based derives from the study of Mahncke et al. (2006). These authors trained 60-87 years-old adults for 8-10 weeks (5 days every week) with exercises based on auditory perception, attention and memory. An improvement was registered in the experimental group in memory assessed with standard testing at retest. However, the analyses did not incorporate the scores from experimental and control group together in a single test (for example a mixed-design ANOVA), but two separate paired t tests were carried out, limiting thus the interpretation of the findings. The same statistical flaw is detectable in a Japanese study based on a different paradigm. A training programme was designed from the hypothesis that the continuous activation of areas involved in high-order cognitive processes would have benefits on more basic cognitive functions. Uchida & Kawashima (2008) designed a protocol of sentence-reading and primary-school arithmetic exercises normally tapping high-order cognitive processes and computational brain regions. Seventy-year-old-and-older adults were enrolled and trained daily for 6 months. An improvement was recorded in the DSMT and 2 subtests of the FAB (lexical flexibility and conflicting instructions). Also a Spanish (Buiza et al., 2008; Yanguas et al., 2006) and a Pennsylvanian (Eckroth-Bucher & Siberski, 2009) study did not combine their groups together in the analyses, and therefore the positive findings associated to their treatment should be interpreted with caution, or, safely, as a simple trend.

The website www.mybraintrainer.com is an online cognitive stimulation programme that has been recently tested by Simpson and her team (2012) on a small sample of 50-to-70 years old adults: twenty-one days of daily web-based exercises tapping a multitude of cognitive skills improved processing speed only, with no impact on aspects of working memory, learning and executive functions.

Finally, three studies used Nintendo Wii-implemented videogames and found conflicting results: one month of game training with Big Brain Academy did not result in any significant cognitive improvement, in comparison to one month of reading (Ackerman et al., 2010). One month of Brain Age resulted instead in a selective improvement in executive functions and speed of processing in comparison to one month of Tetris (Nouchi et al., 2012). Finally, six weeks of Brain Training resulted in a selective improvement in the digit span backwards compared to a no-contact group (McDougall & House, 2012).

The material by Mahncke et al. (2006) was the only package of exercises based on neural evidence. To date, no study has investigated the impact of DMN-based CA on validated
cognitive tools. We therefore analysed the change in neuropsychological performance in our sample of treated and untreated adults. We also analysed the trend of performance on the training material, and its correlation with the change in resting-state brain function.

5.1.3.2 METHODS
The same participants recruited in Experiments 4 and 5 took part in the correlational analysis. Thirty participants took part in the behavioural part of this study, 22 in the experimental group and 8 in the control group. Their characteristics are reported in Table 5.2 (Page 155) and Table 5.11.

Table 5.11 Demographics of participants.

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Education</th>
<th>Gender (M/F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete Sample</td>
<td>68.2 (1.5)</td>
<td>10.0 (0.8)</td>
<td>13/17</td>
</tr>
<tr>
<td>Experimental Group</td>
<td>67.9 (1.7)</td>
<td>10.4 (1.1)</td>
<td>10/12</td>
</tr>
<tr>
<td>Control Group</td>
<td>69.7 (3.5)</td>
<td>8.7 (0.9)</td>
<td>3/5</td>
</tr>
</tbody>
</table>

For the sole behavioural analyses, there were seven additional healthy participants enrolled with the same inclusion and exclusion criteria (6 in the experimental group plus 1 control). These participants could not undergo scanning, and for this reason did not take part in previous two experiments. The programme of CA is described in Section 5.1.1.2.2.1.

5.1.3.2.1 NEUROPSYCHOLOGICAL ASSESSMENT
A comprehensive battery assessing various aspects of cognition was put together with particular focus on the aspects most targeted by normal and pathological ageing: the MMSE (Folstein, Folstein, & McHugh, 1975) was selected as a screening measure of overall cognition; the Digit Span Forwards (Orsini et al., 1987) and Backwards served as measures of verbal short term and working memory; the spatial span was measured with the Corsi Test (Orsini et al., 1987); the same equipment served for assessing visual learning, through the Spatial Supraspan Test (Spinnler & Tognoni, 1987); the Rey-Osterrieth Complex Figure (Caffarra, Vezzadini, Dieci, Zonato, & Venneri, 2002a) was used as a measure of visuoconstructive skills (copy) and visual long-term memory (10-minute delayed recall); verbal memory was assessed with the Prose Memory Test (Novelli et al., 1986a) and the Paired Associates Test (Novelli et al., 1986a); lexical and semantic skills were tested with multiple tools: a Confrontational Picture Naming

---

15 The cognitive assessment was performed by Dr Francesca Bevilacqua, Dr Jessica Rigon, and Dr Cristina Pilosio, I.R.C.C.S San Camillo, Venice, Italy.
Test, the Letter Fluency Test (Novelli et al., 1986b) and the Category Fluency Test (Novelli et al., 1986b); the Digit Cancellation Test (Spinnler & Tognoni, 1987) was used as a measure of visual search and speed of processing; a component of executive functioning was assessed with a paper version of the Stroop Test (Caffarra, Vezzadini, Dieci, Zonato, & Venneri, 2002b); the Token Test (Spinnler & Tognoni, 1987) was chosen as a measure of comprehension and receptive language; the Similarities Test from the WAIS-R battery and the Raven Progressive Matrices (Carlesimo et al., 1996) were respectively chosen as verbal and non-verbal measures of abstract reasoning. All tests consisted of the Italian version. Raw scores on tests were corrected for age and level of education (some also for gender) for clinical purposes, using various sources of normative data except for the Digit Span and the Confrontational Naming, for which no normative data exist for the Italian population. Corrected scores were finally converted into equivalent scores (0 to 4), which were used for diagnostic purposes. The measures of cognitive skills in the experimental section of our study consisted instead of the raw scores.

The interval between baseline and retest assessment ranged between 20 and 105 days (mean=51.4; SEM=3.8) in the experimental group. In this group the treatment started from 1 to 70 days after the baseline evaluation (mean=18.8; SEM=3.5) and the retest evaluation started between 1 and 59 days after the end of the treatment (mean=9.6; SEM=2.8). In the control group the test-retest interval was on average 69.1 days (SEM=10.6; range=35-97) and did not differ from the experimental group (p>0.05).

5.1.3.3 PROCEDURE

The procedure was the same as described in Sections 5.1.1.3 and 5.1.2.3. The only difference was the cognitive assessment, which was performed with the same time modalities as the brain scans.

To evaluate any change in performance over the twenty sessions, average scores were computed for the first and last batch of sessions; since no rapid RT were specifically asked of the participants, number of errors was analysed for all exercises belonging to the Semantic Retrieval and Logical Reasoning categories. Both errors and RT were instead analysed for tasks belonging to the Speed of Processing category. No analysis was run for the remaining categories. Mean values were computed averaging the first and the last batch of sessions; the first and the last session were discarded and averages were created from respectively sessions 2-3-4-5 and 16-17-18-19. Paired t tests were run to compare early and late scores; composite indices were also created by averaging early and late average scores of each stimulation
category. Assumptions for analysis of covariance were not checked, given the exploratory
design of this analysis.

Resting-state functional scans were acquired and analysed as described in Section 5.1.1.2.2.2 and 5.1.1.2.2.3. For the correlational analysis we used the contrast images produced during first level analysis, and we ran a multiple regression adjusting for age and education. We assumed that behavioural changes in the training material were statistically more sensitive to change than a score in a neuropsychological test. We identified only one category of tasks that did not have any increase in difficulty over the sessions: Speed of Processing. Within this category we identified tasks whose performance displayed a statistically significant improvement from the beginning to the end of the training and limited the correlation to the change in performance shown on these tasks.

The neuropsychological assessment was performed at baseline and retest. The battery of cognitive tests was usually completed in 1 to 3 sessions. Scores were analysed via paired t tests, and group comparison was carried out running 2X2 mixed-design ANOVAs. A p value of 0.05 was chosen as threshold of significance, and a Bonferroni correction for multiple comparisons was applied (p<0.0028).

5.1.3.4 RESULTS

Significant increases in errors were found in both the semantic and the logical composite. As for the semantic subtasks, more mistakes were made in late sessions for “Rule the Odd One Out 1” and “Rule the Odd One Out 2”; as for the logical reasoning subtasks, more mistakes were made in “Sequence Completion 2”, “Sentence Completion” and “Scene Completion” (Table 5.12).

Errors increased because sessions must be considered as a function of difficulty level. Trials difficulty increased from early to late sessions, and for this reason it is difficult to interpret any change in performance as an improvement. The only tasks in which challenge was the same from session 1 to session 20 were subtasks of the Speed of Processing category. For this theoretical justification, we chose to use these scores in the correlational analysis.
Table 5.12 Accuracy and RT for trials of each of the subtasks.

<table>
<thead>
<tr>
<th>EXERCISE</th>
<th>EARLY SCORE</th>
<th>LATE SCORE</th>
<th>p</th>
<th>TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule the Odd One Out 1 - ERRORS</td>
<td>0.14</td>
<td>0.34</td>
<td>&lt; 0.05</td>
<td></td>
</tr>
<tr>
<td>Rule the Odd One Out 2 - ERRORS</td>
<td>0.69</td>
<td>2.40</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>Double Category - ERRORS</td>
<td>1.94</td>
<td>1.81</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Semantic Inhibition - ERRORS</td>
<td>1.65</td>
<td>1.33</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Semantic Composite - ERRORS</td>
<td>1.11</td>
<td>1.47</td>
<td>&lt; 0.005</td>
<td></td>
</tr>
<tr>
<td>Sequence Completion 1 - ERRORS</td>
<td>2.86</td>
<td>2.75</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Sequence Completion 2 - ERRORS</td>
<td>0.93</td>
<td>1.40</td>
<td>&lt; 0.01</td>
<td></td>
</tr>
<tr>
<td>Sentence Completion - ERRORS</td>
<td>0.88</td>
<td>1.56</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>Scene Completion - ERRORS</td>
<td>0.42</td>
<td>2.43</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>Reasoning Composite - ERRORS</td>
<td>1.27</td>
<td>2.03</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>Respond to the A - TIME (ms)</td>
<td>157.20</td>
<td>151.08</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Respond to the Square - TIME (ms)</td>
<td>160.64</td>
<td>149.14</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Respond to A &amp; B - TIME (ms)</td>
<td>384.31</td>
<td>366.09</td>
<td>&lt; 0.01</td>
<td></td>
</tr>
<tr>
<td>Respond to Blue &amp; Red Squares - TIME (ms)</td>
<td>377.01</td>
<td>358.16</td>
<td>&lt; 0.05</td>
<td></td>
</tr>
<tr>
<td>Respond to A &amp; B - ERRORS</td>
<td>4.40</td>
<td>3.17</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Respond to Blue &amp; Red Squares - ERRORS</td>
<td>4.23</td>
<td>2.77</td>
<td>n.s.</td>
<td></td>
</tr>
</tbody>
</table>

Within this category we looked at the performance of the treated group, together with the group of treated MCI individuals described in experiments 7, 8 and 9. We decided to conjoin the two samples because we wanted to use the same behavioural variable in the correlational analysis of both groups. The only variables that displayed a significant improvement from early to late sessions in the two groups unified was the error rate in the "Respond to A & B" and "Respond to Blue & Red Squares" tasks (both ps<0.05). We computed an average between the two tasks for both timepoints and calculated a change score.

\[ \text{Change Score} = [(A & B_{\text{early}} + \text{Blue & Red Squares}_{\text{early}})/2] - [(A & B_{\text{late}} + \text{Blue & Red Squares}_{\text{late}})/2] \]

We obtained positive scores for decreases in errors and negative scores for increases in errors. In this way interpretation of correlational analysis is simpler, as positive scores indicate improvements. The correlational analysis on the MCI individuals is described in Experiment 9.
Fig. 5.11 Positive correlation between change in training performance and increase in activation in the healthy subgroup.

Table 5.13 Areas of increased activity significantly correlated to change in training performance.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/ Right</th>
<th>BA</th>
<th>N° voxels</th>
<th>Z value at Local Maximum</th>
<th>x</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle Temporal Gyrus</td>
<td>L</td>
<td>WM</td>
<td>120</td>
<td>4.29</td>
<td>-50</td>
<td>-64</td>
<td>10</td>
</tr>
<tr>
<td>Superior Temporal Gyrus</td>
<td>L</td>
<td>22 GM</td>
<td>110</td>
<td>4.28</td>
<td>-65</td>
<td>-40</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>13 GM</td>
<td>3.60</td>
<td>-57</td>
<td>-42</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>WM</td>
<td>3.60</td>
<td>-51</td>
<td>-44</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Corpus Callosum</td>
<td>L</td>
<td>WM</td>
<td>52</td>
<td>3.83</td>
<td>-6</td>
<td>-44</td>
<td>11</td>
</tr>
<tr>
<td>Frontal Sub-Gyrus</td>
<td>R</td>
<td>WM</td>
<td>16</td>
<td>3.71</td>
<td>36</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Cerebellum - Culmen</td>
<td>R</td>
<td>GM</td>
<td>28</td>
<td>3.61</td>
<td>14</td>
<td>-37</td>
<td>-7</td>
</tr>
<tr>
<td>Parahippocampal Gyrus</td>
<td>R</td>
<td>28 GM</td>
<td>20</td>
<td>3.10</td>
<td>18</td>
<td>-14</td>
<td>-16</td>
</tr>
</tbody>
</table>

Improvement in performance was correlated with augmented activity at rest mainly in the left temporal lobe and the right parahippocampal gyrus, with the involvement of cerebellum, corpus callosum and a subcortical component of the frontal lobe (Table 5.13; Figure 5.11).

The associated decrease was similarly detected in temporal areas but in the right hemisphere, together with right frontal and occipital clusters (Table 5.14; Figure 5.12).
Table 5.14 Areas of decreased activity significantly correlated to change in training performance.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/Right</th>
<th>BA</th>
<th>N° voxels</th>
<th>Z value at Local Maximum</th>
<th>Tailarach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle Temporal Gyrus</td>
<td>R</td>
<td>WM</td>
<td>68</td>
<td>4.12</td>
<td>63 -27 -5</td>
</tr>
<tr>
<td>Superior Temporal Gyrus</td>
<td>R</td>
<td>GM</td>
<td>3.34</td>
<td>57 -27 0</td>
<td></td>
</tr>
<tr>
<td>Middle Temporal Gyrus</td>
<td>R</td>
<td>WM</td>
<td>18</td>
<td>3.63</td>
<td>53 -1 -22</td>
</tr>
<tr>
<td>Lingual Gyrus</td>
<td>R</td>
<td>18 GM</td>
<td>20</td>
<td>3.58</td>
<td>2 -78 2</td>
</tr>
<tr>
<td>Cuneus</td>
<td>R</td>
<td>18 GM</td>
<td>25</td>
<td>3.56</td>
<td>18 -80 27</td>
</tr>
<tr>
<td>Frontal Sub-Gyral</td>
<td>R</td>
<td>WM</td>
<td>23</td>
<td>3.29</td>
<td>20 22 22</td>
</tr>
<tr>
<td>Lingual Gyrus</td>
<td>R</td>
<td>WM</td>
<td>24</td>
<td>3.13</td>
<td>12 -62 1</td>
</tr>
</tbody>
</table>

The paired-sample t test revealed no statistical difference in the majority of the tested domains (Table 5.15). The only test that displayed a significant improvement was the Prose Memory Test, whereas the phonemic fluency and both copy and recall of the Rey Figure were marginally significant as they did not survive adjustment for multiple comparisons. More generally, no test displayed a trend towards worse performance, indicating that the programme was generally beneficial in the experimental group.

Fig. 5.12 Positive correlation between change in training performance and decrease in activation in the healthy subgroup.
Only 2 mixed-design ANOVAs resulted to be significant. In the MMSE score a significant but not meaningful interaction was found: the experimental group remained stable, whereas the control group improved. In the Raven Matrices the interaction indicated an improvement in the experimental group and a decrease in the control group. These two interactions were significant with $p<0.05$ but did not survive correction for multiple comparisons.

**5.1.3.5 DISCUSSION**

Decrease in accuracy in the training material was found between the first and the last days of CA. This negative change can be positively interpreted as a counterbalance between increased difficulty of the trials and a behavioural improvement. The only tasks in which an improvement was seen are those with no increasing level of difficulty. This indicates that our training material was constantly challenging throughout the 20 sessions, as CA material should be demanding enough to be effective (Jaeggi et al., 2011; Zelinski & Reyes, 2009).
The overall picture indicates that the improvement in cognitive efficiency displayed from the beginning to the end of the training was associated with a pattern of redistribution of resting-state activation. This shift was mainly observed from right temporal areas (in which activity decreased) to left temporal areas. Additional occipital activity was lost, and increased activity was recorded in cerebellum, parahippocampal gyrus and corpus callosum. These findings give further support to the functional results described in Experiment 4 (Table 5.2), where the left parahippocampal gyrus, the cerebellum and the right temporal cortex were already reported as showing increase in resting-state brain function. We found that this increase was also correlated with improvement in cognitive performance in the training material. Even though the score in the RT-tasks was chosen for the correlation because of its properties, the association reflected also the intensive work with the other tasks of semantics and logical reasoning. We chose to correlate the neural data to the training material based on the assumption that raw scores on cognitive tests are less sensitive to minimal changes in cognition. Nonetheless we analysed the data relative to the cognitive performance of our healthy participants comparing their baseline and their retest performances. Applying the Bonferroni Correction for 18 comparisons we expected that no improvement could survive a really small \( p \) value. We found instead that improvement in long-term memory was still significant, and that in long-term visual memory it approached significance. These findings suggest that our CA programme was particularly beneficial for explicit memory, which is the cognitive domain that is most sensitive to the effects of normal ageing and AD. A general trend towards improvement was found for all tests, and the improvement in letter fluency and in the copy of the Rey Figure was significant, indicating that our programme may also trigger a benefit in other aspects of cognition. However, the change in verbal memory was not significant in the analysis that also comprised the control group. From the mixed-design ANOVA, a selective benefit was found in global cognition and visual abstract reasoning, but only with an uncorrected \( p \) value. In summary, as suggested at the beginning of this chapter, scores on neuropsychological tests are not optimal for statistical analysis. Nonetheless there is moderate evidence that our programme of CA triggered an improvement in verbal long-term memory and visual abstract reasoning, and weak evidence that it triggered an improvement in other cognitive domains.
5.2 A NETWORK-BASED PROGRAMME OF COGNITIVE TRAINING ON A SAMPLE OF MCI

5.2.1 EXPERIMENT 7 - EFFECT ON BRAIN FUNCTION

5.2.1.1 INTRODUCTION
The applicability and effectiveness of CA programmes have rarely been studied in relation to brain function in samples of MCI. The priority has always been given to improving cognition and variables related to daily life, and for this reason the literature is limited. Studying CA with a focus on neuroplasticity would mean managing to identify structural mechanisms that could be induced via training and lead to significant improvement, measurable through analysis of neurocognitive variables. Hampstead and colleagues (2011) piloted a small CA programme on a sample of 6 amnestic multidomain MCI patients (with no control group). Five sessions of face-name association training were administered over a period of 2 weeks. The first and the fifth session were recorded in the fMRI scanner and compared, together with the analysis of effective connectivity, a parameter similar to functional connectivity, but aimed at detecting a causal relationship between the activities of two separate hubs. The authors interpreted the wide-spread increase in task-related activation as enhancement of areas belonging to the DMN, with enhanced connectivity between the middle temporal gyrus and precuneus/PCC and other parietal areas. Rosen et al. (2011) treated a small sample of MCI patients (of various subtypes, some of them treated also with AD-related medication) with the same programme on which the IMPACT trial was based (see Section 5.1.3.1). An intense regime was chosen, with 24 sessions of individually-tailored difficulty and gradually-increased duration. Efficacy was assessed as a change in activation during performance in a verbal encoding task. Increases were reported in the hippocampus, suggesting that this area, although being extremely susceptible to AD pathology, still retains capacity for plasticity in the MCI stage. Results from a study investigating the efficacy of strategy learning for improving memory performance brought further evidence for plasticity mechanisms in a sample of 15 MCI patients. Encoding- and retrieval-related resources benefitted from training and a rearrangement of activation in a complex pattern of areas was observed after only 12 hours of intervention (Belleville et al., 2011). In conclusion there is some evidence suggesting that people in the MCI stage still retain capacity for neuroplasticity, but more and specific evidence is needed from bigger samples and with whole-brain methodologies of investigations.
We wanted to administer our CA programme to a sample of patients at high risk of developing AD who already show objective cognitive problems. We expected an activation change that could be interpreted as a shift towards a normal pattern of resting state. Testing this in a sample diagnosed with MCI would permit the clarification of whether this expected effect can be stimulated in the presence of a possible neurodegenerative disorder.

5.2.1.2 METHODS
Apart from the participants, the complete methodology was the same as in Experiment 4.

5.2.1.2.1 PARTICIPANTS
Geographical characteristics of the sample were as with participants recruited for Experiments 1-3. Potential MCI individuals were enrolled in multiple ways. In a considerable proportion of cases they had been referred for neurological assessment by their G.P., and during examination were asked whether they were keen on taking part in a project on the benefits of computer exercises on brain function. In other instances a word-of-mouth procedure was used and local G.P.s were requested to identify suitable candidates. A neuropsychological battery (Section 5.1.3.2.1) was administered and diagnosis of MCI was made following the classic criteria (Petersen, 2004). Briefly, participants had to be non-demented, independent in their ADL, and having single or multiple domains of subjective and objective cognitive impairment.

Further exclusion criteria were set as follows: significant pharmacological treatments with psychotropic medication, ChEI, memantine, drugs for research purposes or with toxic effects to internal organs, presence of significant disease at clinical level, a previous history of transient ischemic attacks, a diagnosis of vascular dementia, a baseline structural MRI revealing a different diagnostic pattern from those expected in a MCI diagnosis, presence/diagnosis of uncontrolled seizures, peptic ulcer, cardiovascular disease, sick sinus syndrome, neuropathy with conduction difficulties, significant disabilities, abnormal baseline levels of folates, vitamin B12 or thyroid stimulating hormone.

No control group of untreated MCI patients was recruited.

Visual inspection of realignment parameters suggested that further analysis was needed for the scans of two participants; in particular we wanted to avoid that test-retest differences in translation parameters altered the estimation of the analysis models. For this reason we compared the parameters associated to the baseline scan with those associated to the retest scan. Values did not distribute normally, and non-parametric tests were run, correcting for multiple comparisons (p =0.0167): none of the axes revealed a significant difference for the
first of the two cases, whereas two of the three axes were associated with a significant $p$ value for the other one ($x: p<0.001$; $y: p<0.05$; $z: p<0.01$). For this reason, the latter patient was excluded from the rest of the analyses. The participants included in this experiment are described in Table 5.16 and 5.17.

**Table 5.16 Demographic characteristics of the sample.**

<table>
<thead>
<tr>
<th>Age</th>
<th>Education</th>
<th>Gender (M/F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCI Sample</td>
<td>74.1(2.2)</td>
<td>8.1 (1.6)</td>
</tr>
</tbody>
</table>

**Table 5.17 Cognitive characteristics of the sample at baseline.**

<table>
<thead>
<tr>
<th>Test</th>
<th>N</th>
<th>Corrected Score</th>
<th>Cut-Off</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE †</td>
<td>9</td>
<td>25.56</td>
<td>N/A</td>
</tr>
<tr>
<td>Raven Matrices</td>
<td>9</td>
<td>27.46</td>
<td>18.96</td>
</tr>
<tr>
<td>Digit Cancellation</td>
<td>9</td>
<td>45.75</td>
<td>30</td>
</tr>
<tr>
<td>Stroop Test - Time</td>
<td>8</td>
<td>27.14</td>
<td>36.92</td>
</tr>
<tr>
<td>Stroop Test - Errors</td>
<td>8</td>
<td>6.08</td>
<td>4.24 *</td>
</tr>
<tr>
<td>Phonemic Fluency</td>
<td>9</td>
<td>32.56</td>
<td>16</td>
</tr>
<tr>
<td>Semantic Fluency</td>
<td>9</td>
<td>33.22</td>
<td>24</td>
</tr>
<tr>
<td>Token Test</td>
<td>9</td>
<td>32.86</td>
<td>26.25</td>
</tr>
<tr>
<td>Confrontational Naming †</td>
<td>9</td>
<td>18.00</td>
<td>N/A</td>
</tr>
<tr>
<td>Similarities ††</td>
<td>9</td>
<td>10.44</td>
<td>N/A</td>
</tr>
<tr>
<td>Digit Span Forward</td>
<td>9</td>
<td>5.17</td>
<td>3.5</td>
</tr>
<tr>
<td>Digit Span Backwards †</td>
<td>9</td>
<td>3.00</td>
<td>N/A</td>
</tr>
<tr>
<td>Paired Associates Learning</td>
<td>9</td>
<td>11.33</td>
<td>6</td>
</tr>
<tr>
<td>Short Story Recall</td>
<td>9</td>
<td>15.07</td>
<td>15.76 *</td>
</tr>
<tr>
<td>Visuospatial Span</td>
<td>9</td>
<td>4.69</td>
<td>3.25</td>
</tr>
<tr>
<td>Visuospatial Supraspan</td>
<td>6</td>
<td>10.24</td>
<td>5.5</td>
</tr>
<tr>
<td>Rey Figure - Copy</td>
<td>9</td>
<td>27.67</td>
<td>28.87 *</td>
</tr>
<tr>
<td>Rey Figure - Recall</td>
<td>9</td>
<td>12.03</td>
<td>9.46</td>
</tr>
</tbody>
</table>

* Performance under cut-off score
† Raw score is indicated
†† WAIS-R score is indicated

**5.2.1.3 PROCEDURE**

As the training material was challenging, a 15-minute test with a sample of the CA material was performed at discretion of the neurologist responsible for the recruitment before the participants made their final decision on whether to take part in the study, and to ensure that they were able to do the training exercise competently. The rest of the procedure was as in
Experiment 4. As no control group was recruited, data analysis was carried out with a paired $t$ test.

**5.2.1.4 RESULTS**

Baseline scans were obtained 1-45 days prior to the beginning of the treatment (mean=16.9; SEM=5.4), and retest scans within 4 days after the end of the treatment. The test-retest interval was 22-74 days (mean=42.6; SEM=5.7).

The CA programme triggered a mild overall increase in brain function, with significant clusters located in the precuneus bilaterally, and left frontal and temporal areas (Table 5.18; Figure 5.13). Decreases were found in a bigger number of clusters, and most of these were located in the frontal lobe: middle, superior and medial frontal gyri bilaterally and the ACC on the left side. A scattered decrease was also found in the temporal lobe and in the right Inferior parietal lobule (Table 5.19; Figure 5.14). This picture can be interpreted as a considerable overall down-regulation of frontal lobe activity, whose regions are pathologically up-regulated by AD (Jones et al., 2011).

![Fig. 5.13 Increases observed in the MCI subgroup after treatment: only a few clusters were reported to be significant.](image)

**Table 5.18** Areas displaying augmented activation after training.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/Right</th>
<th>BA</th>
<th>N° voxels</th>
<th>Z value at Local Maximum</th>
<th>Tailarach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precuneus</td>
<td>L</td>
<td>WM</td>
<td>41</td>
<td>4.49</td>
<td>-8 -52 53</td>
</tr>
<tr>
<td>Parietal Sub-Gyral</td>
<td>L</td>
<td>WM</td>
<td>3.45</td>
<td>-14 -50 59</td>
<td></td>
</tr>
<tr>
<td>Superior Temporal Gyrus</td>
<td>L</td>
<td>WM</td>
<td>15</td>
<td>3.88</td>
<td>-38 12 -22</td>
</tr>
<tr>
<td>Precuneus</td>
<td>R</td>
<td>WM</td>
<td>15</td>
<td>3.87</td>
<td>18 -52 48</td>
</tr>
<tr>
<td>Inferior Frontal Gyrus</td>
<td>L</td>
<td>GM</td>
<td>33</td>
<td>3.19</td>
<td>-50 21 0</td>
</tr>
</tbody>
</table>
Table 5.19 Areas displaying decreased activation after training

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/ Right</th>
<th>N°</th>
<th>Z value at Local Maximum</th>
<th>Talairach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle Temporal Gyrus</td>
<td>L WM</td>
<td>18</td>
<td>3.97</td>
<td>-42 -63 26</td>
</tr>
<tr>
<td>Medial Frontal Gyrus</td>
<td>L WM</td>
<td>17</td>
<td>3.94</td>
<td>-8 39 40</td>
</tr>
<tr>
<td>Medial Frontal Gyrus</td>
<td>L WM</td>
<td>23</td>
<td>3.81</td>
<td>-6 55 6</td>
</tr>
<tr>
<td>Middle Frontal Gyrus</td>
<td>R WM</td>
<td>19</td>
<td>3.77</td>
<td>44 12 47</td>
</tr>
<tr>
<td>Middle Frontal Gyrus</td>
<td>R  6 GM</td>
<td>59</td>
<td>3.38</td>
<td>6 46 32</td>
</tr>
<tr>
<td>Cuneus</td>
<td>R  9 GM</td>
<td>19</td>
<td>3.66</td>
<td>38 12 53</td>
</tr>
<tr>
<td>Frontal Sub-Gyral</td>
<td>L WM</td>
<td>39</td>
<td>3.57</td>
<td>-32 -8 37</td>
</tr>
<tr>
<td></td>
<td>L WM</td>
<td>41</td>
<td>3.54</td>
<td>42 -36 17</td>
</tr>
<tr>
<td>Inferior Parietal Lobule</td>
<td>R WM</td>
<td>19</td>
<td>3.25</td>
<td>46 -34 25</td>
</tr>
<tr>
<td>Middle Temporal Gyrus</td>
<td>L WM</td>
<td>23</td>
<td>3.48</td>
<td>-42 -42 6</td>
</tr>
<tr>
<td>Medial Frontal Gyrus</td>
<td>R  9 GM</td>
<td>59</td>
<td>3.38</td>
<td>6 46 32</td>
</tr>
<tr>
<td>Medial Frontal Gyrus</td>
<td>L  9 GM</td>
<td>49</td>
<td>3.12</td>
<td>-6 45 15</td>
</tr>
<tr>
<td>Superior Frontal Gyrus</td>
<td>L WM</td>
<td>40</td>
<td>3.25</td>
<td>-20 50 21</td>
</tr>
<tr>
<td>Anterior Cingulate</td>
<td>L  32 GM</td>
<td>17</td>
<td>3.16</td>
<td>-18 45 7</td>
</tr>
<tr>
<td>Frontal Sub-Gyral</td>
<td>L WM</td>
<td>15</td>
<td>3.25</td>
<td>-20 8 42</td>
</tr>
<tr>
<td>Superior Frontal Gyrus</td>
<td>L WM</td>
<td>17</td>
<td>3.11</td>
<td>-18 57 9</td>
</tr>
</tbody>
</table>

Fig. 5.14 Decreased activation in the MCI subgroup: most of the significant clusters were detected in the frontal lobe, with a marginal involvement of parietal, temporal and limbic structures.

5.2.1.5 DISCUSSION

We administered a programme of CA to a small sample of MCI individuals and we observed a change in brain function that was qualitatively different from that found in healthy participants (Section 5.1.1.4). In the pathological group benefits resulted in a limited up-regulation of posterior DMN. A more considerable down-regulation of anterior DMN was instead achieved, indicating that neuroplastic processes were also triggered in this subgroup, resulting in resting-state functional changes in a direction consistent with our hypothesis, but manifesting
themselves in a qualitatively different picture, that is probably dependent on the baseline status of brain function.

Our study brings further evidence to the fact that neuroplasticity is possible at the MCI stage. Our findings lend more support to this hypothesis than previously published studies, as our sample was slightly larger than those of previous studies (Hampstead et al., 2011; Rosen et al., 2011) and we used a resting-state-, rather than a task-associated approach (Belleville et al., 2011). This suggests that DMN-stimulating non-pharmacological treatment protocols may be effective options in the treatment of conditions in which the early cognitive difficulties might be symptom of AD.
5.2.2 EXPERIMENT 8 – EFFECTS ON BRAIN STRUCTURE

5.2.2.1 INTRODUCTION

The evidence we collected in Experiment 7 and published papers (Belleville et al., 2011; Hampstead et al., 2011; Rosen et al., 2011) suggests that it is possible to modify neural parameters during the MCI stage. As suggested by Lövdén (2010), beneficial functional restorations should be accompanied by structural changes. In Experiment 5 we found substantial GM increase in a sample of healthy individuals following our CA programme. We would therefore expect neural changes in the same direction in a sample of MCI. However, we are aware that baseline neural structure may be different in a stage that could potentially reflect the presence of neurodegenerative processes. For this reason we did not expect the same pattern of change that emerged in the healthy subgroup.

As there are no studies that have investigated change in neural density as a function of CA, we maintained an explorative approach without putting forward specific hypotheses.

5.2.2.2 METHODS

Methodology was the same as in Experiment 5 (Section 5.1.2.2). Inclusion and exclusion criteria are described in Section 5.2.1.2. The MCI participant excluded in Experiment 7 was included in this study.

Baseline characteristics of the sample are detailed in Tables 5.17 and 5.20.

Table 5.20 Demographic characteristics of the sample.

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Education</th>
<th>Gender (M/F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCI Sample</td>
<td>73.8</td>
<td>8.2</td>
<td>5/4</td>
</tr>
</tbody>
</table>

5.2.2.3 PROCEDURE

The procedure was the same as that described in Experiments 7.

5.2.2.4 RESULTS

No changes were detected between baseline and retest levels of total GM and WM. As no control group was available, TIV was not extracted as a possible nuisance variable.
GM changes were widespread in the four lobes, with significant decreases in the left precuneal and lingual gyri, and right inferior temporal and superior frontal cortices; additional clusters were found in the cerebellum (Table 5.21; Figure 5.15). No GM increase was found.

**Table 5.21 GM decrease.**

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/ Right</th>
<th>N° voxels</th>
<th>Distance (mm)</th>
<th>Z value at Local Maximum</th>
<th>Tailarach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lingual Gyrus</td>
<td>L</td>
<td>18</td>
<td>1</td>
<td>4.11</td>
<td>-14 -76 -8</td>
</tr>
<tr>
<td>Cerebellum - Culmen</td>
<td>L</td>
<td>86</td>
<td>0</td>
<td>3.74</td>
<td>-24 -51 -19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.54</td>
<td>-32 -48 -23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.05</td>
<td>-30 -56 -22</td>
</tr>
<tr>
<td>Inferior Temporal Gyrus</td>
<td>R 20</td>
<td>22</td>
<td>0</td>
<td>3.67</td>
<td>51 -9 -20</td>
</tr>
<tr>
<td>Inferior Temporal Gyrus</td>
<td>R 20</td>
<td>23</td>
<td>1</td>
<td>3.59</td>
<td>48 -10 -35</td>
</tr>
<tr>
<td>Superior Frontal Gyrus</td>
<td>R 6</td>
<td>18</td>
<td>0</td>
<td>3.42</td>
<td>6 -5 63</td>
</tr>
<tr>
<td>Cerebellum - Declive</td>
<td>L 18</td>
<td>18</td>
<td>0</td>
<td>3.38</td>
<td>-40 -63 -17</td>
</tr>
<tr>
<td>Precuneus</td>
<td>L</td>
<td>31</td>
<td>3</td>
<td>3.35</td>
<td>-18 -67 25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>3.11</td>
<td>-14 -63 18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>2.90</td>
<td>-14 -70 33</td>
</tr>
<tr>
<td>Lingual Gyrus</td>
<td>L</td>
<td>26</td>
<td>0</td>
<td>3.27</td>
<td>-14 -64 0</td>
</tr>
</tbody>
</table>

**Fig. 5.15 GM decrease in the MCI subgroup: a widespread loss was detected in multiple areas with no well-defined pattern and involvement of all lobes.**

Conversely, a considerable set of WM increase was found, especially in the right hemisphere, where augmentation emerged in the medial cingulate, parahippocampal gyrus, occipital regions and a cluster close to brainstem (Table 5.22; Figure 5.16). A bilateral increase in perithalamic WM was also found. A single decrease was found instead in a large cluster in proximity to the right lateral ventricle (Table 5.23; Figure 5.17).
Table 5.22 WM increase.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/Right</th>
<th>N° voxels</th>
<th>Distance (mm)</th>
<th>Z value at Local Maximum</th>
<th>Tailarach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-lobar Extra-Nuclear</td>
<td>L</td>
<td>53</td>
<td>3.84</td>
<td>-12</td>
<td>0 4</td>
</tr>
<tr>
<td>Lingual Gyrus</td>
<td>R</td>
<td>204</td>
<td>3.79</td>
<td>26</td>
<td>-74 0</td>
</tr>
<tr>
<td>Occipital Sub-Gyral</td>
<td>R</td>
<td>0</td>
<td>3.71</td>
<td>30</td>
<td>-64 0</td>
</tr>
<tr>
<td>Middle Occipital Gyrus</td>
<td>R</td>
<td>0</td>
<td>3.45</td>
<td>26</td>
<td>-77 8</td>
</tr>
<tr>
<td>Precentral Gyrus</td>
<td>L</td>
<td>31</td>
<td>3.70</td>
<td>-46</td>
<td>-2 35</td>
</tr>
<tr>
<td>Cingulate Gyrus</td>
<td>R</td>
<td>28</td>
<td>3.61</td>
<td>14</td>
<td>-14 34</td>
</tr>
<tr>
<td>Sub-lobar Extra-Nuclear</td>
<td>R</td>
<td>39</td>
<td>3.51</td>
<td>16</td>
<td>-25 5</td>
</tr>
<tr>
<td>Parahippocampal Gyrus</td>
<td>R</td>
<td>28</td>
<td>3.32</td>
<td>38</td>
<td>-43 -5</td>
</tr>
</tbody>
</table>

Fig. 5.16 WM increase in the MCI subgroup: similarly to the GM decrease, a widespread pattern was found, with involvement of basal, occipital and motor areas.

Fig. 5.17 WM decrease in the MCI subgroup: a single large cluster was found in the dorsal section adjacent to the inferior horn of the left ventricle.
5.2.2.5 DISCUSSION

Our programme of CA triggered structural changes in a sample of MCI. Apart from a single WM loss, the main pattern was a trade-off between GM loss and WM increase. Similar to the findings of Experiment 7, a recalibration of neural resources was observed. In this case, a GM loss was found, in a scattered set of areas. The loss in the precuneus was also observed in the healthy subgroup (Tables 5.6 and 5.9) and appears to be a general response to the training. The rest of these areas do not specifically belong to well-determined networks, and 5 of the 7 clusters were located outside hubs of the posterior DMN. The concurrent increase in WM was mainly located in subcortical and occipital regions. However, the most relevant finding was the increase in parahippocampal and cingulate volume, two areas that are highly involved in regional connectivity. The medial cingulate is a strategic area that connects the ACC and prefrontal cortex to the PCC and precuneus, in other words a connective area between the two components of the DMN. The parahippocampal gyrus is instead directly involved in the WM reduction observed in AD (Villain et al., 2008). It is therefore possible to conclude that the GM-WM trade-off was in the direction of a positive redistribution of volume, and for this reason the CA programme was effective. These are the first findings demonstrating that brain volume changes can be induced in key-areas by intensive CA. It will be necessary to replicate these findings with the addition of a control group.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/ Right</th>
<th>No. voxels</th>
<th>Distance (mm)</th>
<th>Z value at Local Maximum</th>
<th>Talairach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal Sub-Gyral</td>
<td>L</td>
<td>184</td>
<td>0</td>
<td>3.98</td>
<td>-34 -56 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>3.53</td>
<td>-34 -48 10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>2.96</td>
<td>-34 -41 4</td>
</tr>
</tbody>
</table>
5.2.3 EXPERIMENT 9 – EFFECT ON COGNITION

5.2.3.1 INTRODUCTION

As MCI is purely a cognitive concept, many studies have tried to induce training-dependent changes in neuropsychological functions. A first batch of studies was based on learning and application of cognitive strategies. Eight weeks of teaching memory strategies, with complementary exercises for attention and processing speed led to improvement in a face-name association task and in verbal delayed recall (Belleville et al., 2006). Although the intervention was based on strategy teaching, it is significant to note how MCI reported a type of benefit that is not of theoretical support to models of neuroplasticity, but is an immediate response to practical complaints voiced by a population displaying memory problems in everyday life. Similar findings were recently reported by other teams (Hampstead et al., 2012; Moro et al., 2012). An unsuccessful attempt was made by a Swedish team, who recruited 15 amnestic single-domain MCI (with no control group) and treated them for 8 weeks with a cognitive strategy-based training for everyday life goals. No benefit was reported in memory and executive tasks, and just an improvement in processing speed was registered (Londos et al., 2008). A fifth strategy-based study reported positive effects in samples of healthy adults with subjective memory complaints, but statistical methodology was once again based on intra-groups comparisons only (Craik et al., 2007). In a sixth instance, no benefit was evident in a sample of MCI from the ACTIVE study (Unverzagt et al., 2009).

Four studies of computerized CA training were recently reviewed (Faucounau, Wu, Boulay, De Rotrou, & Rigaud, 2010). The oldest one, written by Gunther et al. (2003) describes a pilot programme based on a German software commercialised in 1992. After 14 weeks the sample (not specifically diagnosed with MCI but with “age-associated memory impairment”) improved on various cognitive measures. Unfortunately, no control group was included. In another study, smaller (n=10) samples of AD patients with an MMSE≥22 (on medication) and MCI patients (not on medication) were administered with a modified version of a software originally developed to treat aphasias. A significant (but uncorrected for multiple comparisons) \( p \) value was found for increases in scores on the MMSE, phonemic fluency and TMT-B in AD patients after two 4-week periods of treatment spaced out by a 6-week break, whilst the MCI group exhibited a significant increase in the Rivermead Behavioural Memory Test (Cipriani et al., 2006). The same software was used by Talassi and colleagues (2007) together with occupational and behavioural therapy on a sample of MCI and mildly demented patients (MMSE=15-23). The only positive impact after 3 weeks was an improvement in delayed recall of the Rey Figure, but the reported significance would have not survived adjustments for
multiple comparisons. Finally, in the last of the four studies, Rozzini and co-workers (2007) used a computerised programme (TNP software) originally intended to target the symptoms of Parkinson’s disease, to train MCI individuals who were also medicated with one of the three ChEIs. Four weeks of intensive stimulation constituted a block, and three subsequent blocks were administered with two-month inter-block distance. The mix of ChEI treatment and this programme focusing on various aspects of cognition led to improvement in the short story recall and the Raven matrices compared to the drugs-only and the no-treatment group. One extra paper (Barnes et al., 2009) described the use of the IMPACT training material in a controlled trial with a training regime similar to the original instructions. Positive trends were reported in the wished direction. The authors concluded that larger samples were needed.

As the most common cognitive problem in MCI is memory, most of the computerised and non-computerised designs have been based on memory exercises and have tried to obtain an improvement in the memory domain. The overall pattern of findings appears to be forged by sporadic success (Stott & Spector, 2011), although reviews often report optimistic conclusions (Gates, Sachdev, Fiatarone Singh, & Valenzuela, 2011), and the findings of a meta-analysis of 17 studies taken together appear to indicate that there might be significantly beneficial results in executive functioning, memory and overall cognition (Li et al., 2011). Once again, however, too much variability undermines the theoretical assumptions to interpret these findings. Six of those 17 studies had no control group and experimental groups ranged from 9 to 67 participants.

None of these studies featured a programme of exercises specifically tailored to target the DMN. For this reason, we tested the efficacy of our programme on a small sample of adults diagnosed with MCI. We analysed test scores, performance on the training material and its correlation with changes in resting-state activation.

5.2.3.2 METHODS
The participants that took part in this experiment were the same as in Experiment 8.

The methodology was the same as in Experiment 6.

5.2.3.3 PROCEDURE
Procedures were as described in Experiment 6.

5.2.3.4 RESULTS
The test-retest interval was in average 49.9 days long (SEM=6.3; range=20-79).
The number of errors increased in the subtests “Rule the Odd One Out” 1 and 2, and in “Scene Completion” (Table 5.24). Both composite indices of errors were significantly higher in the late batch of sessions. A trend for increase in errors number was detected for all other tasks of semantic access and logical reasoning, indicating MCI patients performed similarly to the healthy individuals described in Experiment 7.

The formula detailed in Section 5.1.3.4 was used on the errors made in the Speed of Processing tasks.

Table 5.24 Accuracy and RT in the training material.

<table>
<thead>
<tr>
<th>EXERCISE</th>
<th>EARLY SCORE</th>
<th>LATE SCORE</th>
<th>p</th>
<th>TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule the Odd One Out 1 - ERRORS</td>
<td>0.47</td>
<td>1.06</td>
<td>&lt; 0.05</td>
<td></td>
</tr>
<tr>
<td>Rule the Odd One Out 2 - ERRORS</td>
<td>2.28</td>
<td>4.25</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>Double Category - ERRORS</td>
<td>3.22</td>
<td>3.53</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Semantic Inhibition - ERRORS</td>
<td>2.89</td>
<td>2.94</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Semantic Composite - ERRORS</td>
<td>2.22</td>
<td>2.94</td>
<td>&lt; 0.05</td>
<td></td>
</tr>
<tr>
<td>Sequence Completion 1 - ERRORS</td>
<td>4.08</td>
<td>4.83</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Sequence Completion 2 - ERRORS</td>
<td>2.39</td>
<td>3.11</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Sentence Completion - ERRORS</td>
<td>1.58</td>
<td>2.33</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Scene Completion - ERRORS</td>
<td>1.17</td>
<td>4.14</td>
<td>&lt; 0.005</td>
<td></td>
</tr>
<tr>
<td>Reasoning Composite - ERRORS</td>
<td>2.31</td>
<td>3.60</td>
<td>&lt; 0.005</td>
<td></td>
</tr>
<tr>
<td>Respond to the A - TIME (ms)</td>
<td>161.35</td>
<td>175.72</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Respond to the Square - TIME (ms)</td>
<td>202.71</td>
<td>161.33</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Respond to A &amp; B - TIME (ms)</td>
<td>440.94</td>
<td>469.28</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Respond to Blue &amp; Red Squares - TIME (ms)</td>
<td>436.96</td>
<td>470.21</td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>Respond to A &amp; B - ERRORS</td>
<td>11.59</td>
<td>6.43</td>
<td>&lt; 0.05</td>
<td></td>
</tr>
<tr>
<td>Respond to Blue &amp; Red Squares - ERRORS</td>
<td>10.61</td>
<td>5.11</td>
<td>n.s.</td>
<td></td>
</tr>
</tbody>
</table>

Improvement in training material was associated to increased activation in a large set of areas (Table 5.25; Figure 5.18) widespread to the whole brain. The association with decreases in activation was significant only in a cerebellar and a frontal cluster (Table 5.26; Figure 5.19)
**Table 5.25** Areas of increased activity significantly correlated with change in training performance

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/Right</th>
<th>BA</th>
<th>N° voxels</th>
<th>Z value at Local Maximum</th>
<th>Tailarach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precuneus</td>
<td>R</td>
<td>7</td>
<td>GM 27</td>
<td>4.32</td>
<td>14 -59 54</td>
</tr>
<tr>
<td>Sub-lobar Extra-Nuclear</td>
<td>R</td>
<td>WM</td>
<td>139</td>
<td>4.26</td>
<td>32 -56 9</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>WM</td>
<td>4.10</td>
<td>24 -42 14</td>
<td></td>
</tr>
<tr>
<td>Temporal Lobe Sub-Gyral</td>
<td>R</td>
<td>WM</td>
<td>3.81</td>
<td>36 -41 3</td>
<td></td>
</tr>
<tr>
<td>Posterior Cingulate</td>
<td>L</td>
<td>WM</td>
<td>57</td>
<td>4.24</td>
<td>-28 -58 12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WM</td>
<td>2.91</td>
<td>-22 -60 7</td>
<td></td>
</tr>
<tr>
<td>Precentral Gyrus</td>
<td>R</td>
<td>WM</td>
<td>22</td>
<td>4.01</td>
<td>28 -23 52</td>
</tr>
<tr>
<td>Frontal Sub-Gyral</td>
<td>L</td>
<td>6</td>
<td>GM 36</td>
<td>3.88</td>
<td>-26 -3 56</td>
</tr>
<tr>
<td>Middle Frontal Gyrus</td>
<td>L</td>
<td>6</td>
<td>GM 3.54</td>
<td>-32 -3 51</td>
<td></td>
</tr>
<tr>
<td>Cerebellum - Declive</td>
<td>R</td>
<td>GM</td>
<td>24</td>
<td>3.87</td>
<td>42 -63 -21</td>
</tr>
<tr>
<td>Cerebellum - Culmen</td>
<td>L</td>
<td>GM</td>
<td>23</td>
<td>3.83</td>
<td>-10 -30 -8</td>
</tr>
<tr>
<td>Insula</td>
<td>R</td>
<td>GM</td>
<td>35</td>
<td>3.74</td>
<td>44 2 -2</td>
</tr>
<tr>
<td>Superior Temporal Gyrus</td>
<td>R</td>
<td>WM</td>
<td>3.08</td>
<td>50 4 -7</td>
<td></td>
</tr>
<tr>
<td>Middle Occipital Gyrus</td>
<td>L</td>
<td>18</td>
<td>GM 34</td>
<td>3.71</td>
<td>-8 -96 14</td>
</tr>
<tr>
<td>Cuneus</td>
<td>L</td>
<td>WM</td>
<td>3.28</td>
<td>-14 -95 8</td>
<td></td>
</tr>
<tr>
<td>Paracentral Lobule</td>
<td>R</td>
<td>5</td>
<td>GM 18</td>
<td>3.64</td>
<td>10 -42 56</td>
</tr>
<tr>
<td>Inferior Occipital Gyrus</td>
<td>L</td>
<td>19</td>
<td>GM 31</td>
<td>3.55</td>
<td>-36 -80 -3</td>
</tr>
<tr>
<td>Middle Occipital Gyrus</td>
<td>L</td>
<td>18</td>
<td>GM 37</td>
<td>3.53</td>
<td>-28 -85 5</td>
</tr>
<tr>
<td>Cuneus</td>
<td>L</td>
<td>WM</td>
<td>3.20</td>
<td>-24 -77 10</td>
<td></td>
</tr>
<tr>
<td>Lingual Gyrus</td>
<td>L</td>
<td>WM</td>
<td>2.70</td>
<td>-16 -79 5</td>
<td></td>
</tr>
<tr>
<td>Precentral Gyrus</td>
<td>R</td>
<td>6</td>
<td>GM 17</td>
<td>3.50</td>
<td>36 -2 42</td>
</tr>
<tr>
<td>Superior Temporal Gyrus</td>
<td>L</td>
<td>GM</td>
<td>109</td>
<td>3.42</td>
<td>-44 13 -7</td>
</tr>
<tr>
<td>Insula</td>
<td>L</td>
<td>GM</td>
<td>3.39</td>
<td>-44 2 4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>GM</td>
<td>3.13</td>
<td>-40 -2 4</td>
<td></td>
</tr>
<tr>
<td>Precentral Gyrus</td>
<td>L</td>
<td>WM</td>
<td>30</td>
<td>3.40</td>
<td>-51 -11 10</td>
</tr>
<tr>
<td>Superior Temporal Gyrus</td>
<td>L</td>
<td>WM</td>
<td>3.31</td>
<td>-51 -8 -1</td>
<td></td>
</tr>
<tr>
<td>Cuneus</td>
<td>R</td>
<td>WM</td>
<td>25</td>
<td>3.40</td>
<td>24 -87 11</td>
</tr>
<tr>
<td>Middle Occipital Gyrus</td>
<td>R</td>
<td>WM</td>
<td>2.95</td>
<td>28 -83 5</td>
<td></td>
</tr>
<tr>
<td>Insula</td>
<td>L</td>
<td>GM</td>
<td>20</td>
<td>3.21</td>
<td>-36 -23 13</td>
</tr>
<tr>
<td>Middle Frontal Gyrus</td>
<td>R</td>
<td>WM</td>
<td>17</td>
<td>3.16</td>
<td>22 1 59</td>
</tr>
</tbody>
</table>
Fig. 5.18 Areas of increased activation that display positive association with increase in training performance: a large set of regions were reported: part of these areas are illustrated here (respectively, from left to right, the insula, white matter clusters adjacent to the ventricles, and two frontal and cerebellar clusters).

Table 5.26 Areas of decreased activity significantly correlated to change in training performance.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/Right</th>
<th>BA</th>
<th>N° voxels</th>
<th>Z value at Local Maximum</th>
<th>Tailarach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebellum - Anterior Lobe</td>
<td>R</td>
<td>GM</td>
<td>28</td>
<td>3.91</td>
<td>16 -44 -30</td>
</tr>
<tr>
<td>Medial Frontal Gyrus</td>
<td>R</td>
<td>9</td>
<td>26</td>
<td>3.08</td>
<td>4 48 26</td>
</tr>
</tbody>
</table>

Fig. 5.19 Correlation between decrease in resting-state activation and change in training performance: only two clusters were significant, located in the frontal and cerebellar cortices.

The baseline assessment was carried out 1-44 days prior to the beginning of the treatment (mean=20.3; SEM=5.5) and the retest assessment started within 8 days after the end of the treatment. Based on this the time interval between baseline and retest evaluation was 49.9 days (SEM=6.3; range=20-79).
Baseline and retest performance on cognitive tests revealed one single improvement in the MMSE, which, however, did not survive the correction for multiple comparisons (Table 5.27).

### Table 5.27 Cognitive performance before and after training

<table>
<thead>
<tr>
<th>Test</th>
<th>Baseline N</th>
<th>Retest N</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE</td>
<td>25.6 (0.8)</td>
<td>27.3 (0.9)*</td>
</tr>
<tr>
<td>Raven Matrices</td>
<td>24.1 (2.4)</td>
<td>23.6 (3.2)</td>
</tr>
<tr>
<td>Digit Cancellation</td>
<td>44.1 (3.3)</td>
<td>48.4 (2.5)</td>
</tr>
<tr>
<td>Stroop Test - Time</td>
<td>37.6 (7.3)</td>
<td>34.0 (8.3)</td>
</tr>
<tr>
<td>Stroop Test - Errors</td>
<td>7.6 (2.8)</td>
<td>3.2 (2.2)</td>
</tr>
<tr>
<td>Phonemic Fluency</td>
<td>27.2 (2.8)</td>
<td>26.8 (2.2)</td>
</tr>
<tr>
<td>Semantic Fluency</td>
<td>27.0 (2.4)</td>
<td>27.9 (2.1)</td>
</tr>
<tr>
<td>Token Test</td>
<td>32.7 (1.0)</td>
<td>31.8 (1.3)</td>
</tr>
<tr>
<td>Similarities</td>
<td>16.7 (1.8)</td>
<td>15.9 (1.9)</td>
</tr>
<tr>
<td>Confrontalional Naming</td>
<td>18.0 (0.5)</td>
<td>19.0 (0.4)</td>
</tr>
<tr>
<td>Digit Span Forward</td>
<td>4.9 (0.3)</td>
<td>5.2 (0.4)</td>
</tr>
<tr>
<td>Digit Span Backwards</td>
<td>3.0 (0.3)</td>
<td>2.9 (0.3)</td>
</tr>
<tr>
<td>Paired Associates Learning</td>
<td>9.4 (0.9)</td>
<td>9.4 (1.3)</td>
</tr>
<tr>
<td>Short Story Recall</td>
<td>14.2 (2.6)</td>
<td>18.3 (1.7)</td>
</tr>
<tr>
<td>Visuospatial Span</td>
<td>4.3 (0.2)</td>
<td>4.2 (0.2)</td>
</tr>
<tr>
<td>Visuospatial Supraspan</td>
<td>9.3 (3.2)</td>
<td>7.9 (3.2)</td>
</tr>
<tr>
<td>Rey Figure - Copy</td>
<td>26.0 (2.9)</td>
<td>31.0 (1.5)</td>
</tr>
<tr>
<td>Rey Figure - Recall</td>
<td>7.8 (1.6)</td>
<td>9.0 (1.8)</td>
</tr>
</tbody>
</table>

*Significant with p<0.05

### 5.2.3.5 DISCUSSION

In our small group of MCI, the training proved very challenging, as an increase in the number of errors was observed for all training tasks with increasing difficulty. A decrease was instead visible in the Speed of Processing tasks with fixed difficulty, and this change was associated with a large widespread increase in resting-state activity. This change was less specific than that described in healthy participants (Experiment 6), suggesting that MCI patients need a substantial amount of functional rearrangement to obtain changes in behavioural variables.

The performance in the cognitive tests indicated an improvement in long-term memory measures similar to the healthy group, and in all cognitive domains, but only for the MMSE this improvement (uncorrected) was significant. This brings further evidence to the idea that MCI patients have limited capacity for cognitive plasticity compared to cognitively unimpaired adults, and this spared capacity is hard to detect using purely cognitive measures.
5.3 EXPERIMENT 10 - THE ROLE OF THE APOE GENE AS MEDIATOR OF BENEFIT

5.3.1 INTRODUCTION

The $\varepsilon_4$ isoform of the ApoE gene has been studied in cross-sectional designs to understand whether carriers display differences in brain structure or brain function compared to non-carriers. In MCI patients it was associated to lower hippocampal (Fleisher et al., 2005) and amygdalar (den Heijer et al., 2002) volume, and VBM differences were also found in neocortical areas (Venneri, McGeown, et al., 2011). In healthy individuals genotype-dependent differences have been reported in studies that used VBM (Wishart et al., 2006), DTI (Persson et al., 2006), and cortical thickness (Espeseth et al., 2008). Furthermore, the ApoE $\varepsilon_4$ allele has an impact on resting-state brain function (Reiman et al., 2001; Reiman et al., 1996; Reiman et al., 2005). If baseline differences exist between carriers and non-carriers, it could be speculated that carriers and non-carriers may respond differently to a stimulation targeting the nervous system. This has been demonstrated with a pharmacological stimulation (Marchant et al., 2010), but the role of the ApoE gene has not been systematically studied in the literature on CA. One of the main reasons is due to sample size, as small samples do not permit the achievement of comparable groups of carriers and non-carriers. In addition, the impact of the ApoE alleles is of secondary importance in this case, because the priority has been to test the effectiveness of CA. The presence of the $\varepsilon_4$ allele has been sometimes treated as an adjusting factor (Foubert-Samier et al., 2012; Treiber et al., 2011; Wang et al., 2002; Wilson et al., 2002), and seldom as a predictor. Niti et al. (2008) found that the lowest tertile of self-reported engagement in leisure activity was associated with worse MMSE scores one year later in healthy seniors. Furthermore, ApoE status interacted with amount of CA: $\varepsilon_4$ amplified the distance between tertiles, with mid-tertile and highest-tertile carriers displaying significantly lower odds to decline compared to non-carriers. Although this is just a single result, it suggests that it would be important to further investigate the role of the presence of the $\varepsilon_4$ allele in this area of research, especially for clinical relevance (Corder & Caskey, 2009). The best candidate to investigate the impact of the $\varepsilon_4$ isoform is an experimental trial. For this reason we compared the change in brain structure and function between carriers and non-carriers. As eventually the number of carriers was not sufficiently large, we decided not to investigate the impact on cognitive functions.
5.3.2 METHODS

Blood samples were taken once per participant, at any time during the study, for the analysis of the genotype. Genotyping was carried out through real-time PCR, following the methodology described by Calero and her team (2009). The rest of the methodology was as described in Experiments 4-5-7-8.

5.3.3 PROCEDURE

The distribution of the allelic pairs is illustrated in Table 5.28.

Table 5.28 Proportion of ApoE isoforms in our sample.

<table>
<thead>
<tr>
<th>GENOTYPE</th>
<th>ALL</th>
<th>MCI</th>
<th>HEALTHY</th>
</tr>
</thead>
<tbody>
<tr>
<td>ε2ε2</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>ε2ε3</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>ε3ε3</td>
<td>17</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>ε2ε4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ε3ε4</td>
<td>7</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>ε4ε4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

According to the epidemiological data (sample consisting of 2000 individuals) reviewed by Corbo & Scacchi (1999) the proportion of ε2, ε3, and ε4 alleles in Italy is respectively 0.06, 0.849 and 0.091 (p_{ε2}+p_{ε3}+p_{ε4}=1). We compared these values with our proportions. Pearson’s χ² was significant (p < 0.001) indicating that our data differed from the expected proportions. The main contribution to this significance was the proportion of ε2ε2 individuals (who were however excluded from the analyses):

\[
\text{Partial } \chi^2 \epsilon_{2e2} = (0.09-3)^2/0.09 = 94.09; \quad \text{Total } \chi^2 = 96.16
\]

16 The analyses were carried out by Dr. Annamaria Valletungo, I.R.C.C.S San Camillo, Venice, Italy.
To investigate the effect of ε₄ allele, we gathered the structural scans and the functional contrast images of all 5 ε₄ carriers, and selected 5 ε₄ non-carriers that could match the group by diagnosis, age and education. One-way ANOVA revealed no between-group difference in age, education and TIV (all \( p > 0.05 \)); both groups featured 3 individuals with MCI and 2 healthy controls.

A group-by-timepoint mixed-design ANOVA was run on the structural images, and two-sample t tests were run on the retest-to-baseline contrast functional images. Since the two groups were matched, no covariate was inputted.

The rest of the procedure was as described in Experiments 5-6.

5.3.4 RESULTS

The ε₄ allele had little impact on the change in regional GM and WM volume. Non-carriers displayed a selective increase in GM only in the left cerebellum and in the right superior frontal gyrus compared to carriers (Table 5.29; Figure 5.20). Carriers displayed instead a GM increase in two cerebellar clusters and in the left fusiform gyrus (Table 5.30; Figure 5.21), and a WM increase in the left insula (Table 5.31; Figure 5.22) compared to non-carriers.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/Right</th>
<th>BA</th>
<th>N° voxels</th>
<th>Distance (mm)</th>
<th>Z value at Local Maximum</th>
<th>Tailarach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebellum - Culmen</td>
<td>L</td>
<td>41</td>
<td>0</td>
<td>4.04</td>
<td>-14</td>
<td>-44 -21</td>
</tr>
<tr>
<td>Superior Frontal Gyrus</td>
<td>R</td>
<td>29</td>
<td>1</td>
<td>3.28</td>
<td>38</td>
<td>16 47</td>
</tr>
</tbody>
</table>

Fig. 5.20 Non-carriers displayed GM increase in two cerebellar and frontal clusters, compared to carriers.
Table 5.30 GM increase in carriers compared to non-carriers.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/Right</th>
<th>BA</th>
<th>Nº voxels</th>
<th>Distance (mm)</th>
<th>Z value at Local Maximum</th>
<th>Tailarach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebellum - Culmen</td>
<td>L</td>
<td>44</td>
<td>0</td>
<td>3.67</td>
<td>-16</td>
<td>-36 -13</td>
</tr>
<tr>
<td>Fusiform Gyrus</td>
<td>L</td>
<td>28</td>
<td>0</td>
<td>3.59</td>
<td>-57</td>
<td>-5 -23</td>
</tr>
<tr>
<td>Cerebellum - Declive</td>
<td>R</td>
<td>33</td>
<td>0</td>
<td>3.06</td>
<td>28</td>
<td>-51 -13</td>
</tr>
</tbody>
</table>

Fig. 5.21 Areas of GM increase in carriers compared to non-carriers.

Table 5.31 WM increase in carriers compared to non-carriers.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/Right</th>
<th>Nº voxels</th>
<th>Distance (mm)</th>
<th>Z value at Local Maximum</th>
<th>Tailarach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insula</td>
<td>L</td>
<td>29</td>
<td>0</td>
<td>3.35</td>
<td>-30 3 15</td>
</tr>
</tbody>
</table>

Fig. 5.22 Only an insular WM cluster of difference was found between carriers and non-carriers.

A larger number of differences in brain function were reported. Non-carriers obtained a bigger increase in activation in key areas of both DMN components, with significant clusters in right
precuneus, left inferior parietal lobule, right ACC and other frontal regions. Other non-DMN areas displayed also some differences (Table 5.32; Figure 5.23).

Table 5.32 Areas of training-dependent increased activation displayed in non-carriers, compared to carriers.

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/Right</th>
<th>BA</th>
<th>No. of voxels</th>
<th>Z value at Local Maximum</th>
<th>x</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frontal Sub-Gyral</td>
<td>L</td>
<td>WM</td>
<td>37</td>
<td>4.23</td>
<td>-28</td>
<td>27</td>
<td>5</td>
</tr>
<tr>
<td>Paracentral Lobule</td>
<td>R</td>
<td>6</td>
<td>GM 69</td>
<td>3.71</td>
<td>6</td>
<td>-28</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GM 3.55</td>
<td></td>
<td>2</td>
<td>-36</td>
<td>61</td>
</tr>
<tr>
<td>Frontal Sub-Gyral</td>
<td>L</td>
<td>WM</td>
<td>25</td>
<td>3.66</td>
<td>-18</td>
<td>-9</td>
<td>51</td>
</tr>
<tr>
<td>Medial Frontal Gyrus</td>
<td>R</td>
<td>WM</td>
<td>103</td>
<td>3.61</td>
<td>12</td>
<td>-7</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GM 3.36</td>
<td></td>
<td>2</td>
<td>-3</td>
<td>51</td>
</tr>
<tr>
<td>Precuneus</td>
<td>R</td>
<td>WM</td>
<td>53</td>
<td>3.37</td>
<td>6</td>
<td>-74</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WM 2.88</td>
<td></td>
<td>8</td>
<td>-72</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WM 2.77</td>
<td></td>
<td>10</td>
<td>-70</td>
<td>29</td>
</tr>
<tr>
<td>Caudate Head</td>
<td>R</td>
<td>GM</td>
<td>79</td>
<td>3.30</td>
<td>10</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>Frontal Sub-Gyral</td>
<td>R</td>
<td>WM</td>
<td>3.22</td>
<td></td>
<td>14</td>
<td>21</td>
<td>-10</td>
</tr>
<tr>
<td>Anterior Cingulate</td>
<td>R</td>
<td>WM</td>
<td>3.09</td>
<td></td>
<td>6</td>
<td>23</td>
<td>-5</td>
</tr>
<tr>
<td>Insula</td>
<td>L</td>
<td>GM</td>
<td>29</td>
<td>3.26</td>
<td>-34</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>Precentral Gyrus</td>
<td>L</td>
<td>GM</td>
<td>26</td>
<td>3.21</td>
<td>-36</td>
<td>0</td>
<td>37</td>
</tr>
<tr>
<td>Insula</td>
<td>L</td>
<td>WM</td>
<td>32</td>
<td>3.17</td>
<td>-46</td>
<td>-24</td>
<td>16</td>
</tr>
<tr>
<td>Inferior Parietal Lobule</td>
<td>L</td>
<td>WM</td>
<td>32</td>
<td>3.17</td>
<td>-46</td>
<td>-24</td>
<td>16</td>
</tr>
<tr>
<td>Postcentral Gyrus</td>
<td>L</td>
<td>WM</td>
<td>18</td>
<td>3.16</td>
<td>-59</td>
<td>-22</td>
<td>21</td>
</tr>
</tbody>
</table>

Fig. 5.23 Two-sample t test showing the areas of increased activation in the subgroup of non-carriers compared to the subgroup of carriers. Benefits were higher for non-carriers in a set of midline and neocortical areas.

The analysis of the reverse contrast showing instead the differential benefits of ε4 carriers compared to non-carriers, revealed augmented resting-state activity mostly in areas that are...
not part of the DMN, above all in the cerebellum and in the occipital lobe (Table 5.33; Figure 5.24).

**Table 5.33 Areas of training-dependent increased activation displayed in carriers, compared to non-carriers.**

<table>
<thead>
<tr>
<th>Brain Area</th>
<th>Left/ Right</th>
<th>BA</th>
<th>N* voxels</th>
<th>Z value at Local Maximum</th>
<th>Talairach Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebellum - Culmen</td>
<td>R</td>
<td>GN</td>
<td>155</td>
<td>3.73</td>
<td>5 -44 -22</td>
</tr>
<tr>
<td>Cerebellum - Fastigium</td>
<td>L</td>
<td>GN</td>
<td>3.65</td>
<td>-6 -50 -22</td>
<td></td>
</tr>
<tr>
<td>Cerebellum - Declive</td>
<td>L</td>
<td>GM</td>
<td>407</td>
<td>3.85 -18 -81 -14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.92</td>
<td>-24 -67 -19</td>
<td></td>
</tr>
<tr>
<td>Lingual Gyrus</td>
<td>L</td>
<td>GM</td>
<td>3.82</td>
<td>-20 -92 -5</td>
<td></td>
</tr>
<tr>
<td>Lingual Gyrus</td>
<td>R</td>
<td>WM</td>
<td>21</td>
<td>3.81 20 -76 2</td>
<td></td>
</tr>
<tr>
<td>Cerebellum - Declive</td>
<td>R</td>
<td>GM</td>
<td>67</td>
<td>3.55 -12 -67 -16</td>
<td></td>
</tr>
<tr>
<td>Temporal Sub-Gyral</td>
<td>L</td>
<td>WM</td>
<td>24</td>
<td>3.21 -38 -50 3</td>
<td></td>
</tr>
<tr>
<td>Cerebellum - Declive</td>
<td>R</td>
<td>GM</td>
<td>29</td>
<td>3.13 28 -03 -19</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 5.24 Two-sample t test showing the reverse contrast, with areas in which increased activation was bigger for carriers compared to non-carriers: many cerebellar clusters were detected, plus the bilateral lingual gyrus and a temporal cluster.**

**5.3.5 DISCUSSION**

Our sample was probably too small for a satisfying analysis of ApoE genotype. For this reason we interpret these results as a general trend. However, these results support the idea that genotype may play an important modulation of treatment effects, as non-demented participants homozygote for the ε₃ allele displayed an increased benefit after the end of the stimulation protocol compared to a matched sample of ε₄ carriers.
Structural changes displayed little or no differences in GM and WM between carriers and non-carriers. We interpret this finding concluding that the Ɛ4 isoform does not modulate the structural benefits of a CA programme lasting only 20 sessions.

The two-sample t test gave us a general indication about what the impact of the Ɛ4 allele might be on the effects of our stimulation on brain function. The Ɛ4 isoform was described to be associated with differences in parameters of resting-state BOLD signal in healthy individuals (Filippini, MacIntosh, et al., 2009; Westlye et al., 2011), and functional connectivity of the PCC revealed a decreased pattern of correlation with other hubs, suggesting an allele-dependent down-regulation of the posterior component of the DMN (Machulda et al., 2011). Consistent with these findings, in our small sample the possession of one copy of Ɛ4 allele was associated with a significantly more limited benefit in important posterior hubs of the DMN, particularly in the precuneus, which is anatomically contiguous to the PCC used as a seed in the paper mentioned before. The ApoE genotype has never been investigated as specific predictor of the effects of a CA programme; therefore there is no evidence to compare our findings to. More research is needed, in order to understand whether carrying one copy of the Ɛ4 allele slightly diminishes or minimises drastically the effects of a non-pharmacological treatment. This is relevant in order to further understand whether genotype for the ApoE has to be treated like one of the many risk factors, or rather like a risk factor heavily weighted in the development of AD (Crutcher, 2004). At the moment testing for ApoE genotype is not recommended for diagnostic purposes (Low, Yap, & Brodaty, 2010), but it could be helpful to understand to what extent a targeted form of treatment can improve brain function and cognition as a function of the ApoE genotype. In any case, bigger trials are needed to investigate this point.

In addition, our comparison between carriers and non-carriers did not allow us to understand if the Ɛ4 allele has an effect that is diagnosis-dependant. We balanced the two groups by diagnosis, by comparing two mixed sample of non-demented people; however, it is not clear if the Ɛ4 allele genotype amplifies its effects after the onset of a pathological process by magnifying the burden on plasticity. Allelic differences were found in brain function of young healthy adults (Filippini, MacIntosh, et al., 2009; Reiman et al., 2004), but it also appears likely that a steeper difference can be detected after the possible onset of a pathology the Ɛ4 allele may contribute to trigger. Once again, the inclusion of a qualitatively homogeneous proportion of MCI (all amnestic for instance) would contribute to a theoretically better sample.
5.4 GENERAL DISCUSSION

We tested the effectiveness of our CA programme on brain structure, brain function and cognition of a sample of healthy adults and a sample of MCI. In summary, a GM and brain function increase was found in the healthy subgroup, with a concurrent improvement in long-term memory. In the sample of patients the benefits were generally more limited and qualitatively different, with WM-GM and functional trade-off that did not lead to substantial changes in cognition. In addition, the analysis of the ApoE genotype suggests that the $\varepsilon_4$ isoform modulates prominent training-dependent changes in brain function but not in brain structure.

It is important to pinpoint the differences between-groups in the recalibration of brain activity and associated to improvement in training. Baseline brain function was qualitatively different between the two subgroups, and for this reason less room for neural change in the desired direction was expected in the healthy participants. It is therefore necessary to conceal the apparently contrasting evidence in the MCI subgroup of a bigger increase in parieto-temporal resting-state activation compared to the healthy subgroup on one side, and training performance that never reached the level of that displayed by the healthy subgroup on the other side. Two possible explanations are suggested: a first view would suggest that the recruitment of a bigger amount of neural resources is necessary to observe some behavioural change in the MCI subgroup; this is consistent with the idea of reduced plasticity in the potential early stage of a neurodegenerative disease. Accordingly, a healthy group would obtain a bigger improvement with a limited neural change. A second (but complimentary) possibility is that our cognitive stimulation would not be just about the accuracy with the training material measured in a numeric way, but rather about the simple engagement in cognitive elaboration, regardless of the answers given in the trials. Indeed, most of the baseline-retest differences in the training tasks was not significant, indicating that, in other words, effective CA would be about trying to solve the exercise rather than actually solving it; the resting-state increase in MCI would reflect the effort put in the task, and not the responses given.

Structural changes are associated with biological mechanisms that remain unknown. Longitudinal differences induced by neuroplastic mechanisms can be appreciated by the analysis of cognitive and brain function. However, as already stated, a structural change is the requirement for a long-lasting change in function (Lovden et al., 2010). Normal ageing results in GM and WM loss (Ge et al., 2002), and it is likely that the mechanisms responsible for these non-pathological changes are the same responsible for any test-retest differences due to a
specific treatment. As it is well-known in neuroscience, WM is formed by myelinated axons, whereas GM is composed by non-myelinated axons, shorter neurites, synaptic complexes, glial cells (that take up a considerable proportion of brain volume) and neuronal somas; increases/decreases in neural matter have to be related to mechanisms that modify parameters of these cellular components. Myelinated axons are long structures that are capable to connect neurons whose somata are located in distant areas, thus appearing markedly important for neural circuits. For this reason it is possible to speculate that an increase in WM would be related to an increase in physical characteristics of the axons, detectable using structural neuroimaging. DTI is much more sensitive than VBM in detecting minimal WM changes due to experimental manipulation, but its output is interpreted as a function of diffusivity. We found GM increases in the healthy subgroup and WM increase in the MCI subgroup in areas susceptible to AD pathology. The first areas hit by AD are those which develop myelin sheaths in latest development (Braak & Braak, 1997). It could be speculated that MCI brains having early-stage AD pathology recuperate this parameter thanks to training, whereas healthy adults (who do not have anything to recuperate) experience an increase in GM. Ameliorations in WM structures have already been read as increased axonal myelination (Takeuchi et al., 2010) and we could interpret our findings in the MCI subgroup as better connections between distant areas. A link between abnormal myelin breakdown and decreased brain metabolism has been already suggested (Bartzokis et al., 2006), and likewise, we can infer that one month of network-based training improved axonal health in two key connection areas normally affected in AD. The decrease of GM observed in the same subgroup can be similarly interpreted as a non-negative event because it did not selectively involve areas primarily affected by AD pathology, but was more widespread (and the decrease in the precuneus was equally found in the healthy subgroup). A decrease in GM has already been reported as positive effect of CA in other papers (Takeuchi, Taki, Hashizume, et al., 2011; Takeuchi, Taki, Sassa, et al., 2011). GM decreases are normally associated to negative events (e.g. pathologies), but exercise-induced reductions have been interpreted as functionally beneficial. The aforementioned Japanese team interpreted the structural loss as training-dependent elimination of unnecessary synapses, and, as it is about GM, it would be a selective elimination of short-distance connections between adjacent or neighbouring neurons. The loss of this type of connections would reflect a decrease of local connectivity, a parameter that is not a specific AD biomarker. It might be boldly speculated that early neurodegenerative processes causing a reduction of intrinsic connectivity would also increase local connectivity as a compensating phenomenon; the partial restoration of long connections would therefore result in a degeneration of short connections that are no longer required.
It has been suggested that the $\varepsilon_4$ allele may have a different impact on myelination than the $\varepsilon_3$ counterpart (Bartzokis et al., 2006). If this is true, and myelination can be promoted by CA (Takeuchi et al., 2010), then any other structural mechanism associated with neuroplastic changes in brain function might potentially be influenced by the ApoE genotype. Unfortunately these mechanisms have not been fully understood and therefore an exploratory approach must be assumed in the study of the gene.

Most of the findings emerged from Experiments 4-10 can be explained as a function of our experimental expectations. However, additional regional changes have been reported, which were not predicted by our initial hypothesis. An interesting and not specifically expected finding that recurs across the experiments is the involvement of cerebellar structures. This is particularly visible in terms of volumetric exercise-dependent increases in the healthy experimental group, decreases in the MCI group, and both structural and functional differences between carriers and non-carriers of the $\varepsilon_4$ allele. Apart from motor control, brain-activation paradigms and brain-lesion case studies have revealed that the cerebellum is involved in many cognitive domains (Paquier & Marien, 2005). Furthermore, the study of functional connectivity of its sub-regions has determined that the cerebellum takes part in many functional networks (Stoodley, 2012). We did not perform a specific cerebellar subdivision in lobules as we did for cerebral areas via the Talairach client, but we simply inspected the maps of neural differences visually. In many instances the cerebellar clusters overlapped with areas functionally connected with prefrontal regions (O'Reilly, Beckmann, Tomassini, Ramnani, & Johansen-Berg, 2010), suggesting that structural changes in this region could be associated to a meaningful change in a cerebro-cerebellar functional circuitry.

Similarly, $\varepsilon_4$ carriers may have displayed a differential benefit from the training possibly because of baseline differences, or because of differences in the response to treatment. Among the various cognitive aspects the cerebellum is involved in, it is likely that the tasks based on language abilities and semantic processing may have played a crucial role in relation to the changes displayed by this structure, as there is a large body of evidence suggesting a role of structures of the cerebellum in a range of semantic skills (Venneri et al., 2008).

Our study has some potential shortcomings: first, although we included a no-contact control group, we did not include an extra condition in which our structured exercises could have been compared to unstructured training. This was previously done for changes in neuropsychological tests only (Jelcic et al., 2012) but has not been tested in relation to brain function. The lack of a training-based control group may represent an important bias in our findings (Thomas & Baker, 2012), but it is worth considering that in the literature no study has yet reported a specific impact of CA on a specific set of areas in an exploratory design, finding
no involvement of any single frontal cluster. We acknowledge the weakness but, given the pronounced directionality of the findings, we also believe that the support to our hypothesis is strong.

Second, we included a group of MCI individuals who displayed an inhomogeneous pattern of cognitive impairment. It is therefore normal to expect that part of the variability observed in the data might be due to this aspect. This is cause of concern if we want to interpret the results as a possible application of the training in prodromal AD. MCI subtype does not generally predict the type of dementia the individual may convert to (Hussain, 2007), but very often an amnestic MCI condition hides a diagnosis of probable AD. For this reason the ideal sample to test our programme would be a group of amnestic MCI individuals or a group on whom the AD criteria identified a probable prodromal AD (Dubois et al., 2007).

Third, it was impossible for the tester administering the exercises to be blind to the diagnostic group. Blindness to group membership is important to obtain pure unbiased results, but in case of participant-tester interactions, it is usually very common to figure out who is cognitively impaired. Nevertheless, this bias is not expected to act neither on neural data, nor on the neuropsychological assessment, since independent testers were in charge of each assessment.

It is not clear to what extent it would be possible to regain connectivity in a condition like AD that normally features loss of synapses and connections. MCI represents an ideal in-between phase in which awareness of possible pathology is already present, but retained plasticity is sufficient to observe structural and functional improvements in terms of enhanced connectivity. However, it is not true that AD is a disconnection syndrome in which this pruning is observed homogeneously in the brain. If structural connectivity is linked to functional connectivity, evidence of pathological up-regulation of anterior DMN would imply that structural connectivity would be similarly abnormal in the frontal lobe. Therefore, using a logic argument, structural changes are still possible in AD. A CA programme should not aim to generally restore, but instead redirect connectivity. Based on this, the ultimate aim of cognitive stimulation in ageing, MCI and AD would be an optimisation of residual neuroplastic processes.
6 FINAL DISCUSSION

At the moment the management of AD is a clinical challenge, because on one hand there is an enormous body of research describing the pathology from multiple angles, but on the other hand the prognostic perspectives are still really poor. The most invalidating aspect of AD is the behavioural decline (changes in cognition, mood, psyche, independence, and quality of life); any intervention that could sensibly improve these aspects, could be defined as "successful". Nowadays the only approved treatment is a pharmacological approach based on the evidence that enhancing the cholinergic transmission in AD may lead to a symptomatic amelioration. Unfortunately, the rate of responders is not 100%, and side effects of drugs must not be under-estimated. This should stimulate the scientific community to look for other treatment hypotheses and identify a new approach that can offer potential results. Many attempts have been made (Olazaran et al., 2010; Rabipour & Raz, 2012), but not all approaches have been equally successful. It is necessary to identify specific types of stimulation that reasonably have a consistent organic impact on the nervous system.

From a global perspective, non-pharmacological stimulations are different from pharmacological stimulations only as a matter of the initial triggering event. As long as an intervention works as a function of a hypothesised neural mechanism, the nature of the "sparkle" that induces the desired change is technically irrelevant. For this reason, the potential of a non-pharmacological approach is as big as that of a medication. Perhaps, to the ear of a non-expert, treatment for AD is a concept automatically associated to a "wonder-drug" that can miraculously cure the disease, and non-pharmacological interventions might sound like a less-effective approach, maybe just because they are not “chemical enough”: it is not a question of method, but rather of underlying rationale. It is not impossible that in the future a cure might become available, but at present nothing like this is in sight. A considerable proportion of studies have been carried out in vitro or in animal models, and those variables can be analysed in humans only by peripheral, radiological, or post-mortem investigations. These studies have sparked enlightenment on descriptive elements of the disease but rarely have they come up with a possible treatment option. It is necessary to take a wide look at the global literature and identify potential avenues of intervention. In the jungle of axes modified by AD, brain function seems to be the most fruitful. It is directly associated with behavioural variables, but also has a structural counterpart that could be the organic target of a treatment. And, moreover, it represents a useful interpretational interface of the two main non-pharmacological interventions: PA and CA. These two approaches normally judged by common
sense as an effective preventive and therapeutic tool for cognitive decline, are actually supported by some scientific evidence. Unfortunately, the primary interest has often not been describing the results in terms of a strong background hypothesis based on organic variables, but rather interpreting behavioural changes in the name of a more “common-sense-oriented” benefit; brain function offers two interpretative models that afford to understand the mechanism of both CA and PA in a scientifically strong way that is not as distant as a mouse model or a microscope slide.

We wanted to explore the idea that hypotheses based on brain function could account for the effects of specific PA and CA on the brain. We reviewed the literature and identified the theoretical frameworks that represented the ideal viewpoint of both pathology and experimental manipulation of PA and CA. We then defined two convenient paradigms of study to outline the hypotheses that CA and PA would have a benefit on neurocognitive variables in a population subjected to abnormal ageing due to AD pathology. We finally tested the idea that the presence of the $\epsilon_4$ isoform would have an impact on the benefit triggered by CA and PA. We focused our attention on resting-state brain function and identified two theoretical entities: the THT and the DMN are two constructs of resting-state brain function that describe how the brain works in the absence of an explicit mental task. The THT describes a redistribution of resources due to engagement in sufficiently heavy exercise, and the DMN identifies a functional circuit that is affected by certain pathologies. Although these two constructs are far apart from each other, they are both supported by neuroimaging evidence and both can account for long-term changes in brain function, especially if triggered by a focused stimulation.

We designed Experiments 1-3 based on the speculation that the same physiological mechanisms are responsible for the response to both AE and CE. As the evidence suggests a selective impact of PA on the prefrontal cortex, we identified the THT as the construct that may account for the benefits of CE.

In Experiment 1 we tested the THT for the first time in a sample of untrained adults, based on the assumption that the physiology of old and unfit adults and adults at the early stage of a neurodegenerative process would respond in a comparable manner. As untrained adults may display a variable and unpredictable response to exercise, we faced this ample variability by adding to our design a subjective variable that measured the perceived effort of each treadmill session. The results supported the THT, and we speculated that a continuous induction of prefrontal hypo-function would eventually turn into an optimisation of resource consumption.
and a progressive metabolic benefit in the frontal lobe. Once gender was inserted in the model, the findings were only valid for females.

As exercise triggers a stress response, and as a stress response has an impact on brain function and cognition, in Experiment 2 we asked whether stress response played a major role in accounting for the effects of PA. We did not find any significant PA-dependent change in stress response, and we concluded that the effect of PA on the prefrontal cortex is independent from stress.

In Experiment 3 we investigated the impact of the ApoE genotype on the transient prefrontal hypo-function caused by AE. We only found a trend of difference between carriers and non-carriers, and we decided to rerun the analyses separately for the subgroups of males and females, as suggested by the evidence of Experiment 1. In the subgroup of females genotype-associated differences were found in prefrontal-dependent cognition, with carriers and non-carriers displaying different pattern of response to exercise. We interpreted this finding as moderate evidence that the presence of the $\varepsilon_4$ has an impact on prefrontal-dependent cognition even at a young age, and thus differences due to the ApoE genotype must exist in the response to AE, at least in those who are least prone to exercise.

We then created a programme of CA specifically based on the idea of network stimulation and consisting of exercises designed from the early-stage decline of the DMN. We tested the impact of this programme on brain structure, brain function and cognition in a sample of healthy elderly and in a sample of MCI patients.

In Experiment 4 we reported the benefits of our programme on brain function in the subsample of healthy participants. The intra-group and the inter-group comparisons revealed a selective increase of resting-state activity in areas ascribable to the posterior component of the DMN.

In Experiment 5 we found that our programme triggered a selective GM increase in selective and widespread cortical areas, with limited decrease and moderate WM changes.

In Experiment 6 we investigated cognitive changes and we found weak-to-moderate evidence that our programme induced an improvement in long-term memory and abstract reasoning.

A limited increase in brain function was found in Experiment 7 in our sample of MCI, who instead showed a large down-regulation of areas that are part of the anterior component of the DMN. As this component is pathologically up-regulated in early AD, we concluded that
functional ameliorations in MCI are visible in a qualitatively-different, yet, complimentary change in resting-state activity, compared to healthy adults.

The findings of Experiment 8 revealed a CA-dependent well-defined trade-off between unspecific GM loss and WM increase in meaningful areas in our sample of MCI, and in Experiment 9 we found limited evidence of cognitive improvement in global cognition. The correlation between change in performance on the training material and change in brain function revealed a very large pattern of association compared to the correlation found in the healthy subsample described in Experiment 6. This suggested that, in confirmation of the findings of Experiments 7 and 8, MCI patients have limited capacity for neuroplasticity, and they need to induce a large neural restructuring to trigger behavioural changes.

Finally, in Experiment 10 we compared the only 5 $\varepsilon_4$ carriers of our experimental samples to a subgroup of non-carriers matched by age, education, and diagnosis. Little structural differences were found in response to CA, but non-carriers obtained a significantly larger increased activation in meaningful areas, whereas the selective benefit for carriers was limited to regions that are not sensitive to normal or abnormal ageing.

We demonstrated that it is possible to alter parameters relative to the nervous system and modify brain function and cognition with the use of non-pharmacological methods. As PA is believed to induce an enhancement on global brain health in a fashion connected to the theoretical model introduced by the THT, it is possible to conclude that the kernel of the benefit is the frontal lobe. PA can be considered as a beneficial stimulation that could even be defined as a proper form of treatment. In the context of AD, even though evidence is not completely convincing, PA could be used in virtue of its therapeutic advantages. PA does not display a big degree of variability in its operational definition, and it is conceivable that all studies based on oxidative-metabolism PA are actually based on the same experimental manipulation. As for CA, it is likely that every single programme has some degree of uniqueness which makes comparisons difficult. However, the idea of PA as a non-pharmacological treatment against a neurodegenerative disorder suffers from a major shortcoming: the impossibility to choose where specifically the effects will be directed to, inside the brain. CA appears to be more effective in this objective, in fact the evidence reviewed in Chapter 2 has demonstrated that it is possible to modulate it and target specific cerebral domains. When CA programmes have been not designed within a theoretical framework based on neural evidence, there has been too much unstructuredness, and results have been of difficult interpretation, even when the paradigm has been based on brain function but has not focused on the idea of a network. Classifying training programmes into
knowledge-based and process-based training is a first important step to separate CA that leads to structural modification of brain connections from CA that simply taps flexibility and use of alternative strategies; the constant repetition of a particular set of tasks is the core of process-based training, and the idea of remoulding the brain from the prolonged engagement in specific tasks reminds of the structural changes observed after a period of expertise in a particular field or job (Rabipour & Raz, 2012). However, a protocol designed to stimulate activity in one or more single brain regions may not be a sufficiently strong preamble to expect an improvement in brain function. This is another reason why task-associated neuroimaging provides only partial information about the efficiency of a treatment. Apart from its earliest stages where the neural involvement is circumscribed to mediotemporal structures (Braak & Braak, 1991), AD does not affect one brain area only at the time. For this reason it is impossible to obtain a comprehensive picture of the impact of AD on brain function using only one single task-related scan. It is instead necessary to use a paradigm that allows one to take into account activity or metabolism of the whole brain. This does not necessarily mean that any resting-state approach is preferable. Indeed, ROI-based correlational patterns are hypothesis-driven and do not investigate the brain comprehensively. Moreover although seed-based connectivity covers the whole 3D volume of the brain, it only does it in relation to the seed. A seed-based analysis of functional connectivity does not equal an independent component analysis, because the seed might be involved in more networks. For instance, interpreting ACC-seeded resting-state correlation as the salience network (a circuit normally anti-correlated with the DMN) (Machulda et al., 2011) may be an easy-to-manage estimate, but it is not recommended as the ACC takes part in the DMN as well.

In this picture the role of TAU and Aβ appear to be temporarily put on the back burner, as they have not been necessary for the test-retest evaluation of programme efficiency. Even though these molecules are important in the clinical rather than experimental setting, recent studies have investigated the link between specimens and cognitive variables as a function of CA (Landau et al., 2012), PA (Foster et al., 2011), and even cognitive plasticity (Uttner et al., 2011). This might indicate that concentration of proteic markers change dynamically and can be reliably used as test-retest measures in the experimental setting, although more evidence is needed to confirm this.

It has been suggested that the ε4 isoform of the Apolipoprotein E gene elicits an influence on the mechanisms associated with the main proteic biomarkers (Huang et al., 2001; Strittmatter et al., 1993), cholinergic system (Allen et al., 1997), brain function (Trachtenberg, Filippini, Ebmeier, et al., 2012), brain structure (Lind et al., 2006) and cognition (Whitehair et al., 2010). The idea that the ApoE is an independent predictor playing a unique role in AD pathology is
attractive. The idea that this genotypic form modulates the efficacy of an intervention is also attractive. Common methodologies to assay the genotype for the ApoE allow one to obtain a categorical variable that apparently accounts for significant findings with a relatively simple interpretational effort. In all likelihood this is only a simplified picture. There is a large number of genes involved in AD (even just marginally), and any variability is the results of genomic, rather than genetic differences (Fiocco et al., 2008). From this point of view the contribution of ApoE has to be put back in its right perspective, because no striking findings are expected from a single blueprint that does not have any causal role, especially when many of the remaining blueprints have still to be discovered. Even if we found convincing evidence supporting a role of the ApoE gene, this was unfortunately limited because of the sample size. Since data collection is carried out blind to genotype, it is necessary to recruit a sufficiently large sample, as according to most statistics the presence of the $\varepsilon_4$ isoform is only visible in $\frac{1}{4}$ of the individuals.
6.1 STUDY LIMITATIONS

We listed the specific experiment-related shortcomings in the discussion session of each of the studies. However, there are also general limitations in this study that should be highlighted. First, we concluded that PA can be beneficial for frontal metabolism by testing the hypothesis of hypofrontality in a paradigm of AE. Moreover, we also described the effects of stress response and ApoE genotype in relation to acute imbalance of regional metabolism. Our interpretation suggests that the application of this principle may be extended to paradigms of CE, and provide an explanation also to the long-term benefits of exercise. However, this is a conclusion that implies further investigations, and for this reason it can only be considered as a speculation. Our conjecture suggests a detrimental short-term effect and a beneficial long-term effect: in a similar way, the acute stress response and the differences due to genotype might display trends in relation to CE different from those observed in our design. This idea needs to be tested independently. Second, our pattern of findings was derived from a population of healthy adults. As the efficiency of PA would be particularly important for MCI and AD patients, transferring our findings to these diagnostic groups is not automatic, as qualitative differences may exist in the response to manipulation between different groups.

Our programme of cognitive stimulation triggered changes in brain structure and brain function. However, only a portion (albeit a large one) of these differences was predicted and accounted for by the initial hypothesis. There is then part of our pattern of findings that are hypothesis-unspecific and for which we can only make some speculations. Although we designed each exercise on the basis of AD-related deficits in brain function and activation-related knowledge of various cognitive aspects, we did not have perfect awareness of the exact brain regions tapped by every single trial. For this reason, part of our findings remains unspecific. In addition, our programme did not induce any significant improvement in scores on cognitive tests. As a positive change in behavioural variables would be the most desirable consequence of CA, it could be concluded that the pragmatic effects of our programme were insufficient to determine that the training was effective. However, as pointed out in Section 5.1.3, cognitive tests are not appropriate for test-retest statistical comparisons, and for this reason the lack of significant differences does not necessarily mean that no improvement in cognition was promoted.

The main limitation in the study of the effects of the ApoE genotype is a small number of \( \varepsilon_4 \) carriers (this was particularly evident in Experiment 10). Unfortunately there is no remedy to this small proportion, because the study of genetic differences in the brain and cognition depends on the distribution of the various alleles naturally observed in the population.
6.2 SUGGESTIONS FOR FUTURE RESEARCH

- In Experiment 1 we concluded that a progressive enhancement of prefrontal metabolism may be the reason for a long-term programme of PA. This could be directly tested by recruiting young or old adults and measuring resting state brain metabolism at baseline, and after a programme of CE. In addition, it would be interesting to test different diagnostic groups, to see whether similar benefits are observed in conditions with lower levels of neuroplasticity.

- The effect of stress response in Experiment 2 could be studied more successfully in relation to the THT in a more homogeneous sample of athletes.

- The echoplanar scans of Experiments 4 and 7 were analysed in terms of simple brain activation. It would be interesting to reanalyse these data with ICA by looking specifically at the DMN and its changes from baseline and retest. Likewise, it would be informative to look at the follow-up data introduced in Figure 5.1, both the maintenance after three months of inactivity, and the differences between baseline and follow-up, which may indicate the long-term benefits resistant to inactivity.

- A possible better understanding of the real efficacy of our stimulation could be specifically studied by comparing our programme (network-based) to a parallel programme of exercises of similar difficulty targeting the same areas that are part of the network, but singularly, without triggering any synchronisation. This would demonstrate that any differential benefits would not be due to the simple repeated activation of multiple areas, but to their repeated co-activation.

- The "network-based" rationale of cognitive stimulation could be extended to other neurodegenerative disorders that have been described in terms of loss of connectivity. Specific programmes of cognitive stimulation could be designed on the basis of the pictures of connectivity loss observed in each pathology. Even stroke can be described in terms of loss of connectivity (Carter, Shulman, & Corbetta, 2012), and thus even specific rehabilitation for stroke patients could follow our framework.

- The analysis of structural connectivity would also provide important knowledge to the topic.

- The hypotheses of an effect of the ApoE found in Experiments 3 and 10 could be studied as dose-effect response, but this would need a considerably larger sample.
6.3 CONCLUSION

We are going along a transitional stage in the research on AD: in the recent years there has been an increasing interest in the brain from both the scientific community and the non-insiders. A lot of progress has been recently made in the description of the pathology, thanks to technological breakthroughs and application of innovative techniques borrowed from other disciplines: the “standard” MRI and PET scans have been flanked by the addition of advanced statistical procedures or modified histological dyes to take the description of the disease to a new level. Some further steps forward have been made also thanks to the recent network-science based approach, in which anatomical and functional evidence is used to estimate "small-world" parameters that are believed to measure alternative aspects of neural organisation: network science has been proven helpful to detect test-retest changes due to independent factors (Burdette et al., 2010), but it also represents the perfect image of the transitional nature of the period, as it is not clear what parameters emerging from the use of this method are the best ones to characterise a neural system and its changes over time (Bullmore & Sporns, 2009). The transitional stage is also taking shape in the identification of the “Cognitive Stimulation” entity: in the past the idea of “training” or “stimulation” has been used to describe an extremely variable range of programmes that have not just differed in variables such as duration, challenge or frequency, but above all in the theoretical rationale: CA and PA are tightly linked to neurophysiological mechanisms and those have too often been completely bypassed. Some authors have also been interested in assessing the effects of mixed treatments of CA and PA (O’Dwyer et al., 2007; Olazaran et al., 2004; Small, Silverman, et al., 2006), or CA and ChEI (Yesavage et al., 2008), which is believed to trigger a synergic beneficial effect on one side, but on the other side it is very difficult to identify the physiological interactions that may have led to the benefits. Obviously the aim of these trials has been trying to identify an effective protocol of stimulation, which is really needed at the present time, especially in pathological diagnostic groups, and above all in those variables that are associated with daily life. However, for research purposes, proof of effectiveness is not enough; and over the last years more interest has been designated for the underlying mechanisms of brain function and the ideas of reserve and plasticity. The introduction of such concepts has allowed reviews to start distinguishing between knowledge-based and process-based training, between the use of strategies and the extensive repetition of trials, between exercise that manages to trigger oxidative metabolism and exercise that does not. In other words today the picture is much clearer and so is awareness of the specific objectives to reach.
This study supports the idea that non-pharmacological stimulation can be an effective approach in the regulation of parameters of brain function normally de-regulated by healthy ageing or abnormally de-regulated by AD pathology. Our findings bring further support to the idea that oxidation-based PA leads to enhanced metabolism of prefrontal structures. This process can directly exert a positive influence on prefrontal-dependent cognition, and can also be indirectly advantageous as a correlate of compensatory neural computations in support of the metabolic down-regulation of posterior portions of brain areas and brain networks normally observed in AD.

On the other hand, cognitive stimulation is an extremely more heterogeneous concept than physical activity, and our findings suggest that it has to be theory-driven in order to be beneficial. One month of computer-based activities resulted in profound regulation of brain function, accompanied by important changes in brain structure. Based on our framework, cognitive exercises should be designed to tap multiple areas at the same time. This would induce a regulation of interregional connectivity, rather than target single areas. This new conception of cognitive stimulation may be easily translated into the scenario of many other acute and chronic conditions affecting the brain. Rethinking functional loss as a decline in functional connectivity implies that cognitive interventions should be re-designed to regulate the network, rather than simply being symptomatic.

Such evidence suggests that non-pharmacological stimulation may be a valid approach for AD and other-type pathology, and it may be complimentary to the classic protocols of pharmacological treatment. In addition, it can also be considered as a cheap, safe and effective measure of prevention for individuals that are potentially in the preclinical or prodromal stage of pathology.

The analysis of $\varepsilon_4$ carriers and non-carriers suggests that the presence of this risk factor for AD is not simply associated to the descriptive aspect of the pathology (as detailed in Chapter 1), but can also modify the effect of a non-pharmacological treatment. However, no drastic differences were detected, and for this reason this pattern does not suggest enough support for utility of routine ApoE genetic assay (Low et al., 2010).
REFERENCES


impairment and clinical progression to Alzheimer disease. Archives of Neurology, 62(11), 1728-1733.


228


229


230


with healthy control-group members. *Proceedings of the National Academy of Sciences of the United States of America*, 98(6), 3440-3445.


blood flow and middle cerebral artery blood flow velocity during physical exercise.
*Journal of Applied Physiology*, 81(1), 413-418.

. van der Flier, W. M. (2009). Baseline CSF p-tau levels independently predict
doi: Doi 10.1212/Wnl.0b013e3181b879ac

for Sedentary and Active Men and Women. *Arquivos Brasileiros De Cardiologia*, 96(1),
54-59. doi: Doi 10.1590/S0066-782x2010005000155

phosphorylated tau and prediction of progressive mild cognitive impairment.
*Neurology*, 64(7), 1294-1297.

beta 42, Tau and phosphorylated Tau, APOE epsilon 4 allele and MCI type in
progressive MCI. *Neurobiology of Aging*, 28(4), 507-514. doi: DOI
10.1016/j.neurobiolaging.2006.02.001

G. (2010). Apolipoprotein E allele 4 is not a sufficient or a necessary predictor of the
DOI 10.1016/j.eurpsy.2009.02.009

persons with cognitive impairment and dementia: A meta-analysis. *Archives of Physical

01-24). State College, PA 16803, USA: Salimetrics, LLC.

contribute to the pathogenesis of numerous clinical conditions including HSV-1 corneal
10.1016/j.exer.2006.08.001

doi: Doi 10.1016/S0167-8760(03)00080-1

prediction of rapid conversion to Alzheimer's disease in mild cognitive impairment
using regional cerebral blood flow SPECT. *Neuroimage*, 28(4), 1014-1021. doi: DOI
10.1016/j.neuroimage.2005.06.066

The Effects of Physical Activity, Education, and Body Mass Index on the Aging Brain.
*Human Brain Mapping*, 32(9), 1371-1382. doi: Doi 10.1002/Hbm.21113


250


264


Honolulu-Asia Aging Study. *Journals of Gerontology Series a-Biological Sciences and Medical Sciences, 63*(5), 529-535.


