Linking brain and cognitive plasticity in aging:

A longitudinal magnetic resonance imaging study of memory training in middle and late life

Andreas Engvig

Dissertation for the Degree of Ph.D.

Center for the study of human cognition
Department of Psychology
Faculty of Social Sciences
University of Oslo

2011

Series of dissertations submitted to the
Faculty of Social Sciences, University of Oslo
No. 344

ISSN 1504-3991

Cover: Inger Sandved Anfinsen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Unipub, Oslo.
The thesis is produced by Unipub merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

1. Abstract

The goal of the present thesis was to study how memory training is related to structural characteristics of the adult aging brain. We studied two aspects of a novel memory training intervention: 1) is the training accompanied with measureable changes in the brain? (paper I-II), and 2) is training-benefit in adults with memory problems related to structural features of the brain—pre-intervention?

In papers I-II, we undertook a randomized controlled magnetic resonance (MR) imaging study of an intensive two-months memory-training program in healthy adults (mean age 60.3). Assessment of cognitive performance and MR-imaging was performed before and after training. Structural changes in cerebral grey and white matter were assessed using MR-protocols optimized for reliable longitudinal analysis. Following intervention, the training group improved task-specific memory performance. A unique finding is that training was accompanied by regional increases in cerebral cortical thickness and white matter integrity compared with controls. We found significant relationships between the changes in performance and brain structure characteristics, suggesting a link between the two levels of enquiry. The current assessment-interval spanned less than three months, and follow-up studies are needed to conclude on the long-term effects of the present memory training on brain and cognition.

Earlier research indicated positive effects of cognitive intervention for adults with memory concerns, but evidence regarding who might benefit from training was lacking. In paper III, we offered the same training program for memory clinic outpatients with subjective memory impairment, and studied predictors of training benefit. The results showed that training was feasible for this patient group, including high participation rates and low dropout rates. Regional left hippocampal volumes before training were found to predict memory-training benefit. Sub-region analysis suggested that the effects were selective to the left cornu ammonis (CA) sectors CA2/3, and CA4 and dentate gyrus, which are of known importance for episodic memory. The finding implicate that structural imaging could serve useful in future trials evaluating treatment potential and in selecting candidates for intervention.

2. Table of contents

3. Acknowledgements

Several persons made this thesis possible. Professor Kristine Walhovd, my main supervisor together with co-supervisor professor Anders Fjell: The thesis springs entirely out of their scientific curiosity and ability to initiate bold research projects. Their supervision and close follow-up of my progress has been valuable, and for which I am grateful. The social and collegial environment they have created for our research group is unique.

The present research project was funded by the Research council of Norway. I have appreciated salary support from the Kavli Research Center for Ageing and Dementia, led by professor Anette Hylen Ranhoff.

The Medical student research program, Faculty of medicine provided the framework to launch my research career during medical studies. My workplace since 2007 has been the Institute of psychology, currently led by Professor Kjetil Sundet. I am grateful that me as an aspiring medical doctor has been accepted into the Ph.D., -program in psychology.

Lars Tjelta Westlye, co-author on all papers in this thesis has meant a lot for the undertaking of this thesis. I admire his profound ability to set aside time to help, despite being truly busy.

I would like to highlight the contribution of Oslo university hospital, Ullevål, and the Memory clinic in particular. Head of clinic, Dr. Nina Voss Skaane supported our research proposal, and later led patient recruitment from the Memory clinic together with Tor Erling Dahl. Dr. Anne Rita Øksengaard assisted patient recruitment from the memory clinic at Bærum hospital. Øyvind Sundseth was my tutor in neuropsychological assessment and assisted participant recruitment. Torgeir Moberget designed the memory paradigm and aided data gathering. Vivi Agnete Larsen had a role as memory trainer instructor, and aided data acquistion. Dr. Paulina Due-Tønnessen at Rikshospitalet performed neuroradiological evaluation of all participants.

I would like to thank my parents, Kirsti and Jan Gunnar. They have provided the best upbringing one could ask. My mentor, professor emeritus Harald Breivik is a profound source of inspiration. Finally, I would like to thank all the participants.

4. List of papers

The present thesis is based on the three papers listed below. They are referred to in the text by roman numerals.

I. Engvig, A., Fjell, A.M., Westlye, L.T., Moberget, T., Sundseth, Ø., Larsen, V.A., Walhovd, K.B. (2010). Effects of memory training on cortical thickness in the elderly. *NeuroImage* 52(4), 1667-76.

II. Engvig, A., Fjell, A.M., Westlye, L.T., Moberget, T., Sundseth, Ø., Larsen, V.A., Walhovd, K.B. (In press). Memory training impacts short-term changes in aging white matter: A longitudinal diffusion tensor imaging study. *Human Brain Mapping*.

III. Engvig, A., Fjell, A.M., Westlye, L.T., Skaane, N.V., Sundseth, Ø., Walhovd, K.B. Hippocampal subfields volumes correlate with memory training benefit in subjective memory impairment. *Submitted manuscript*.

Amendments:

Since the submission of this dissertation, a revised version of paper III has been published in NeuroImage, volume 61, pages 188-194. The article is available online (doi:10.1016/j.neuroimage.2012.02.072). The revision resulted in minor text alterations only.

5. Acronyms and definitions

ACTIVE = Advanced cognitive training for independent and vital elderly.
AD = Alzheimer´s disease.
CA = Cornu ammonis, i.e., the horn of the Egyptian god Amun. Prefix for distinct anatomical subdivisions of the hippocampal formation.
CNS = Central nervous system.
CVLT-II = California verbal learning test II. A neuropsychological assessment tool for verbal learning and memory.
DG = dentate gyrus. An anatomical subdivision of the hippocampal formation
DTI = Diffusion tensor imaging.
EMQ = Everyday memory questionnaire. A 27-item assessment tool for quantifying subjective memory performance.
FA = Fractional anisotropy.
FMRIB = Oxford Centre for Functional Magnetic Resonance Imaging of the Brain.
FSL = FMRIB Software Library. Collection brain image analysis tools.
GDS = the 30-item Geriatric depression scale.
GLM = General linear model.
IADL = Instrumental activities of daily living.
IQ = Intelligence quotient.
MCI = Mild cognitive impairment.
MD = Mean diffusivity.
MMSE = Mini mental state examination.
MoL = Method of Loci
MP-RAGE = Magnetization prepared rapid gradient echo pulse sequence.
MR = Magnetic resonance.
MRI = Magnetic resonance imaging.
RD = Radial diffusivity.
TBSS = Tract-based spatial statistics. Protocol for voxel-wise analysis of diffusion tensor imaging data.
TFCE = Threshold-free cluster enhancement.
tp = time-point.

5.1 Taxonomy

5.1.1. Brain plasticity

Neurological construct. Herein defined as the ability of the central nervous system to respond to intrinsic and extrinsic stimuli by reorganizing its structure, function and connections (Cramer et al., 2011). I consider brain plasticity as a synonym of the often-used term neuroplasticity—within the brain.

5.1.2. Cognitive plasticity

Psychological construct. Herein defined as the ability to acquire cognitive skills and improve performance following training.

5.1.3. Verbal source memory

Source memory involves judgments about the origin, or source, of information (Johnson, Hashtroudi, & Lindsay, 1993). In the present thesis, verbal source memory denotes recall of the temporal order of previously acquired words.

6. Introduction

6.1. Preface

Aging is associated with changes in brain structure and declining cognitive functioning. Yet, the age trajectories differ among individuals (Buckner 2004). Considerable attention has recently been given to the effects of various forms of cognitive-enrichment in aging (Hertzog, Kramer, Wilson, & Lindenberger, 2008). Memory training may be important in this regard (for a review, see Reichman, Fiocco, & Rose, 2010). At the starting point of this thesis, however, it remained unknown whether improving cognitive performance through memory training is accompanied by measureable changes in the adult and aging brain itself.

The first objective of this thesis was to illuminate this issue by magnetic resonance brain imaging (papers I-II). This study could increase our understanding of the cerebral basis for the suggested role for mental exercise in adulthood and aging.

A second objective was to assess the feasibility of memory training in a clinical setting. Middle-aged and older adults with subjective memory impairment, but with no objective signs of memory loss, made up 27% of memory clinic referrals in 2010 (Brækhus, Ulstein, Wyller, & Engedal, 2011). We offered memory training to a sample of these well functioning, but worried adults referred to Oslo area memory clinics. Here, we assessed whether MR-estimated volumes of known importance for episodic memory predict memory-training benefit. MR volumetry is a putative brain biomarker, but the usefulness of this modality in predicting memory-training outcomes was unknown.

In the following, I review studies from the literature on aging of the brain and cognition. I focus on studies that indicate reductions in specific regions of the brain and certain cognitive abilities, such as episodic memory. Changes in brain and cognition with age seem to be related phenomena. I discuss why brain and cognition might be influenced by the individual´s lifestyle and behavior. The concept of brain plasticity may be important. I emphasize cognitive training and magnetic resonance studies of structural brain plasticity. Next, I focus on a group of persons of particular interest for the study of cognitive interventions. Finally, I present the main objectives of the current thesis.

6.2 Aging and the brain

Aging of the brain has interested researchers for centuries (see, e.g., Kaes, 1907). Anatomical studies of humans were until the late 1900s restricted to autopsy material. Post-mortem investigations revealed many gross and micro-anatomical characteristics of the aging brain, including reductions in brain weight and volume, as well as decline in the bulk and quality of the white matter (Dickstein, Morrison, & Hof, 2009). Confounders, such as cause of death and post mortem tissue alterations limit post-mortem studies. The advent of magnetic resonance imaging (MRI) in the late 1970s revolutionized the field by allowing researchers to study the brains of living individuals. MRI is now a prime research tool in the human neurosciences because it enables 1) reliable, automated, and time-efficient quantification of structural brain characteristics *in vivo* (e.g., Han et al., 2006), 2) more direct comparison with cognitive performance (Cabeza, Nyberg, & Park, 2005), and 3) longitudinal investigations of brain structure, opening for research questions on causation of change (Lindenberger, von Oertzen, Ghisletta, & Hertzog, 2011).

6.2.1. Aging of the grey matter

Similar to the main conclusions from post mortem studies, magnetic resonance imaging studies have demonstrated differences between the brains of young and old individuals (Fotenos, Snyder, Girton, Morris, & Buckner, 2005; Good et al., 2001; Ziegler et al., 2011). Linear age-related thinning of the outermost layer of brain—the cortex—is evident early in adulthood (Fjell, Westlye, et al., 2009; Salat et al., 2004). Deeper, subcortical structures mostly show volumetric reductions with normal aging. In a cross-sectional multi-sample study comprising 883 individuals free of dementia, Walhovd and colleagues (2011) found linear negative effects of age on thalamic volume in all samples. Not all brain structures show consistent or simple linear age effects, however. Age group differences have (Raz, Rodrigue, Head, Kennedy, & Acker, 2004), and have not (Sullivan, Marsh, & Pfefferbaum, 2005), been observed for hippocampal volume in healthy aging. Others have shown that the size of the hippocampus exhibit a non-linear delayed pattern of decline first evident in later parts of the life span (Ziegler, et al., 2011).

Until recently, most knowledge about the aging brain was derived from cross-sectional studies. However, inferring rates of change from age correlations is associated with a number of methodological caveats (Raz & Lindenberger, 2011). Arguably, longitudinal

studies are needed in order to address questions of causality and true intra-individual changes. In a landmark paper, Raz and colleagues (2005) reported a longitudinal study of regional brain volumes in 72 adults scanned at two-time points. The results demonstrated 1) longitudinal volume reduction of the caudate, hippocampi and the association cortices, 2) health factors such as blood pressure affected the rate of shrinkage in some regions, and 3) longitudinal measures of brain shrinkage exceeded most cross-sectional estimates. Note that longitudinal designs have their own limitations, including dropouts and sampling bias at follow-up. In the study by Raz and colleagues, only 23% of the eligible baseline cohort was included in the follow-up analyses. Longitudinal studies of cortical thickness in healthy 60-year olds have suggested *annual* atrophy of 0.5% to 1.0% in most regions of the brain (Fjell, Walhovd, et al., 2009). A longitudinal study of 142 healthy elderly subjects aged 59–90 years, spanning only six months reported mean reductions in volumes of the left and right hippocampi of − 1.02% and − 0.98%, respectively (Murphy et al., 2010). These results suggest more dramatic reductions of this structure in older adults than previously envisioned.

The underlying mechanisms of the MR-observed cerebral cortical atrophy in normal aging are not fully understood. Neuronal loss is seen to a small extent (Pakkenberg & Gundersen, 1997; West, Coleman, Flood, & Troncoso, 1994), but is not thought to play a major role in normal aging (Burke & Barnes, 2006; but see Smith, Rapp, McKay, Roberts, & Tuszynski, 2004). Other explanations come from histological post-mortem analyses indicating that reduction in total neuronal size (Esiri, 2007), reduced number and complexity of dendrites (Duan et al., 2003), reduced spine number and density (Dickstein, et al., 2009), and also myelin disruption (Peters, 2002) accompany normal aging.

6.2.2. Aging of the white matter

Myelin constitutes a major part of white matter. Post mortem studies suggest that cerebral white matter is the brain structure that declines most with age (Marner, Nyengaard, Tang, & Pakkenberg, 2003). Diffusion magnetic resonance imaging yields indirect, but sensitive measures of white matter microstructure (Le Bihan, 2003). By fitting diffusion MR-data to a model—the diffusion tensor—researchers can quantify restricted water movement due to properties of the tissue micro-environment—including axonal membranes, microtubules, and myelin sheats (Beaulieu, 2009).

Resembling data from grey matter investigations, studies with diffusion tensor imaging (DTI) suggest greater white matter deterioration with advancing age (see, e.g., Sullivan & Pfefferbaum, 2006). The most common DTI measures associated with aging include increased mean diffusivity (MD) and decreases in fractional anisotropy (FA) (see Chanraud, Zahr, Sullivan, & Pfefferbaum, 2010, for a review). Regional analyses have revealed stronger adverse effects of age on these measures in anterior as compared to posterior white matter systems (Bennett, Madden, Vaidya, Howard, & Howard, 2010; Davis et al., 2009; Head et al., 2004; Salat et al., 2005; Sullivan & Pfefferbaum, 2006); but exceptions to the pattern exist (Barrick, Charlton, Clark, & Markus, 2010; Giorgio et al., 2010). Nevertheless, increased MD values are thought to represent reduction in overall tissue barriers of the white matter, irrespective of direction; FA decreases are probably related to reduction in the directional coherence of water diffusion along individual fiber tracts.

Myelination has been shown to modulate FA to some degree (Nair et al., 2005; Song et al., 2002; Tyszka, Readhead, Bearer, Pautler, & Jacobs, 2006). A third measure, radial diffusivity (RD) is the average diffusion along the axes of the tensor where bulk water movement is the smallest (lowest diffusion). In one study, age-related decreases in FA were mainly driven by increases in radial diffusivity (RD) (Giorgio, et al., 2010). RD, together with FA and MD, was shown to be a robust predictor of myelin content in post-mortem human brains (Schmierer et al., 2008). This feature of DTI has attracted the attention of cognitive neuroscientists as age-related myelin breakdown has been linked to cognitive decline and Alzheimer´s disease (Bartzokis, 2004). Also, multiple studies have found relationships between age-related variability in DTI-indices and cognitive performance (see Madden et al., 2011, for a recent review). To what extent increased white matter disruption observed with advancing age mediates the observed cognitive differences, or vice versa, remains unknown (Salthouse, 2011).

6.3. Aging and cognition

Multiple properties of cognitive functioning have been investigated in relation to age (Park & Reuter-Lorenz, 2009), again mostly using cross-sectional designs. In these studies older adults are slower, and have less working memory capacity, compared with younger adults. The matter of when cognitive decline begins is one of controversy (Schaie, 2009). As discussed by Nilsson and colleagues (2009), earlier longitudinal studies failed to decline

before 60 years of age. In contrast, the recent Whitehall II study provides longitudinal evidence of decline in multiple functions from age 45 (Singh-Manoux et al., 2012). Others (Finch, 2009; Salthouse, 2009) argue that decrements start earlier, around ages 20 to 30. Precise knowledge about the age at which cognitive decline begins is crucial for future interventions designed to prevent or reverse age-related decrements.

Age influences various functions disproportionately. In an adult life-span study of 345 adults aged 20 to 92, Park and colleagues (2002) reported negative relationships with age for several indices of episodic long-term memory, while semantic memory (e.g., knowledge of facts and vocabulary) remained intact, or even increased with age. The pattern of age-differences also varies depending on the characteristics of the task within cognitive domains (Naveh-Benjamin & Old, 2008). Let us consider individuals´ ability to freely recall words from a previously presented word-list (verbal recall) and their ability to answer yes or no to whether a particular word was on that list (verbal recognition) assessed by the well-studied California verbal learning test II. Normative data for the test indicates substantial effects of age on recall (Delis, Kramer, Kaplan, & Ober, 2000b). Yet, only minor effects of age are seen for the recognition task. Recognition scores on this test remain essentially unchanged—with relatively few false positive errors—into late adulthood (Kramer, Blusewicz, & Preston, 1989).

When provided with a strong cue (e.g., an earlier encountered word), it seems that older adults tend to score as well as younger adults. Conversely, when left with little or no external context or hints, such as in free recall of the CVLT-II, older adults tend to experience more difficulty. The performance discrepancies might stem from the underlying processes thought to sub-serve the various memory tasks. For instance, free recall performance arguably requires more self-initiated processing, strategy use, and cognitive control functions compared to recognition. Research suggests that reductions in self-initiated encoding strategy use partly mediate age-related changes in verbal recall (Verhaeghen & Marcoen, 1994). The brain circuits thought to support episodic memory and other higher-order processes such as strategy use and cognitive control include the medial temporal and frontal cortices (Hedden & Gabrieli, 2004)(Madden). If regional grey matter volumes and white matter integrity of the frontal lobes decline with age—and these regions support strategy and cognitive control processes—how do these phenomena relate to each other?

6.4. Reciprocal relationship between aging of the brain and cognition

Is cognitive aging an inherent negative consequence of aging of its biological substrate, the brain, analogous to reduced vision following aging of the lens and cornea? Or do the two phenomena affect each other reciprocally?

Until recently, joint study of cerebral and cognitive aging was not a distinct research field. The advent of magnetic resonance imaging, however, paved the way for the cognitive neuroscience of aging, a synthesis of disciplines seeking to link cerebral and cognitive aging (Cabeza, et al., 2005). Review-studies within the field indicate that reductions in specific cognitive abilities such as episodic memory are related to neuroanatomical differences, at least statistically; meaning that 25% - 100% of the differences between young and old participants in selected cognitive functions can be explained by group differences in brain structure characteristics (Fjell & Walhovd, 2010). But the question of mediating and causative factors in the age-brain-cognition triangle is highly controversial (Raz & Lindenberger, 2011; Salthouse, 2011). It remains unresolved how and to what extent brain and cognition—matter and mind—impact each other with advancing chronological age. This is due to a number of methodological issues, including cross-sectional design limitations (Lindenberger, et al., 2011).

Still, it is reasonable to believe that part of cognitive decline with aging is indeed mediated by adverse anatomical changes in the brain, constraining upper limits of cognitive functioning. For instance, age atrophy and reduced white matter integrity of the frontal lobes might lead to reductions in cognitive processes operating through frontal lobe circuitry. Here I define neurogenic effects as cognitive reductions mediated by the neurobiological constituents of the brain. Part of the variance in neurogenic effects of age is likely explained by genetics (e.g., differences in sex chromosome and autosomal gene alleles). Nearly two-thirds of genes expressed in the human genome are related to brain function (Petrella, Mattay, & Doraiswamy, 2008). One well-studied genotype in aging research, the ε4 allele of the ApoE gene constitutes a major risk factor for cognitive decline and dementia, and is associated with accelerated age-related cortical thinning in healthy carriers (Espeseth et al., 2008). However, genetics exert differential effects on various brain structures. For instance, the size of the hippocampus has a heritability of only 40% (Sullivan, Pfefferbaum, Swan, & Carmelli, 2001). The reasonable follow up question is:

What about the other 60% ? The answer is undoubtedly complex; it includes gene-environment interactions, and is a topic that lies outside the scope of the present thesis.

Yet, later theoretical work and—to a growing extent—empirical evidence suggest that aging of brain and cognition is a reciprocal process. This account implies that the age trajectories are driven by cognitive, social and environmental aspects of the individual, and biological constraints of age and disease (Hertzog, et al., 2008). Such a framework opens for the possibility that, for instance, failure to engage in certain cognitive processes may cause older adults to suffer greater cerebral atrophy in brain regions supporting these processes (Mahncke, Bronstone, Merzenich, & Aage, 2006). I herein define psychogenic effects on aging as effects on neurocognitive aging that stem from variation in e.g., cognitive activity level and health-maintaining behavior. From the reciprocal viewpoint, psychogenic effects occur in orchestra with neurogenic effects shaping the given age trajectory of the individual.

In this context, psychogenic effects on neurocognitive aging can also be thought of as positive. For instance, Cabeza, Nyberg and Park (2005) argued that older adults who undergo cognitive training might improve cognitive function; the improvement in function could in turn mediate positive biological alterations in relevant brain networks. The scientific rationale for why e.g., improving cognitive functioning over time could impose changes in the brain is in part based on the concept of plasticity (Jäncke, 2009).

6.5. Plasticity

"Plasticity" is derived from the Greek word πλαστ΄οσ (plastos), which means molded. According to the Oxford English Dictionary, being plastic refers to the ability to undergo a change in shape. William James was the first to introduce the term plasticity to the neurosciences in reference to the susceptibility of human behavior to modification. Later, in 1904, Santiago Ramón y Cajal argued that behavioral modifiability must have an anatomical basis in the brain and thus extended the notion of plasticity to the neuronal substrate (Pascual-Leone, Amedi, Fregni, & Merabet, 2005).

6.5.1. Plasticity of the central nervous system—brain plasticity

Decades passed before researchers were able to elucidate the biological underpinnings of the brain´s ability to remodel in response to learning and experience. Then, in 1966, Terje

Lømo at the University of Oslo showed that repeated stimulation of hippocampal neurons led to long-term enhancement of signal transmission (Lømo, 1966), later referred to as long-term potentiation—LTP (Bliss & Lømo, 1973). This is regarded as the first evidence of synaptic plasticity—a suggested mechanism of learning and memory at the cellular level. Later studies showed that synaptic adaptation was coupled to the amino acid glutamate and its respective postsynaptic receptor complexes, AMPA and NMDA—at the molecular level (Luscher, Nicoll, Malenka, & Muller, 2000). Later, Jenkins, Merzenich, and colleagues documented reorganization of entire somatosensory cortical maps after tactile stimulation in owl monkeys (Jenkins, Merzenich, Ochs, Allard, & Guic-Robles, 1990), indicating plasticity at the cortical level (see Buonomano & Merzenich, 1998, for a review).

This is a mere fraction of late 1900-century research illuminating the structural adaptability of the adult central nervous system. However, before the arrival of in vivo imaging methods applicable to humans, scientists had not been able to study this in living persons. A fundamental question remained unsolved: Are differences in the anatomy of the adult human brain predetermined or is the human brain susceptible to plastic change in response to environmental stimulation and learning?

6.5.2. MR-imaging of human brain plasticity

Human neuroscience changed dramatically through the development of MR-imaging methods in the late 1900s. On the structural imaging side, applications such as MR-based morphometry now provided in vivo delineation of temporal changes in brain morphology (May & Gaser, 2006).

Among the earliest studies linking behavioral differences with in vivo estimates of human brain structure were a series of experiments by Schlaug, Jäncke and others. They found that brains of musicians differed notably from that of non-musicians, including alterations in corpus callosal size and gyral folding patterns (Schlaug, Jäncke, Huang, Staiger, & Steinmetz, 1995; Schlaug, Jäncke, Huang, & Steinmetz, 1995). Music-related anatomical differences were more pronounced in musicians who began playing at an early age, pointing to early exposure and genetic disposition as primary drivers of the structural differences. The authors concluded in one the studies:

“it remains uncertain whether gross anatomy may also be susceptible to some postnatal plastic change, such as in response to specific stimulation.” (Schlaug, Jäncke, Huang, & Steinmetz, 1995, pg. 700)

In a seminal paper, Eleanor Maguire and colleagues (2000) showed that the posterior hippocampi of 16 London taxi drivers, who are required to memorize and find all streets in London, were significantly larger compared with controls. This study suggested that the structural difference could be a result of adulthood experience as greater posterior hippocampal volume correlated with longer time spent as a taxi driver. The authors argued:

“it seems that there is a capacity for local plastic change in the structure of the healthy adult human brain in response to environmental demands.” (pg. 4398)

The findings implicated that magnetic resonance imaging could be used to measure experience-related macro-structural adaption in adults, reflected in this case as an increase in grey matter mass. However, new questions arose: Are these changes in fact due to the navigation-related experience of taxi drivers, or are there unknown confounds underlying the observed effects? To answer questions of causality one needs to apply a longitudinal design—studying volunteers involved in a certain form of intervention (e.g., spatial navigation) over time, preferably comparing the cerebral trajectories with those randomly assigned to a control group.

Not long after, Draganski and colleagues published such a study (2004). The authors conducted an experiment where healthy 20-year olds learned to juggle and practiced until they were able juggle three or more balls for at least one minute. Compared with controls, the 12 jugglers exhibited transient grey matter increases in the visual area V5 of the cortex – an area thought to play a major role in the perception of motion. The results were later replicated by the same group (Driemeyer, Boyke, Gaser, Büchel, & May, 2008), and finally by comparable results from an independent research group (Scholz, Klein, Behrens, & Johansen-Berg, 2009). In the latter study, Scholz and his colleagues also found training effects in the underlying cerebral white matter, as revealed by DTI. Conclusively, these studies implicated that 1) experience-related brain plasticity is evident at a macro-scopical level in human adulthood, and 2) MR-imaging holds potential for studying this phenomenon in vivo.

Do middle-aged and older adults exhibit similar MRI-revealed brain plasticity as that shown for younger adults? Animal models suggest age limits of the capacity for adaptive changes (Wagner, Schmoll, Badan, Platt, & Kessler, 2000). The same juggling-experiment was hence performed on a group of 60-year olds. Though plastic changes appeared less pronounced, similar to the young group, grey matter increases related to skill acquisition were found in visual area V5 (Boyke, Driemeyer, Gaser, Buchel, & May, 2008). The authors concluded:

"Given that cortical plasticity of the human brain is, not only on a functional but also on a structural level, preserved even in later years, our data support conclusions about the potential value of exercise for elderly people [...] As people age, they should not do less, but do more to keep and maintain their abilities." (pg. 7034)

Thus, it seemed as training of sensorimotor skills induced brain structural changes in not only the young, but also older adults. In light of these findings, further studies should illuminate how and to what extent other behavioral interventions influence the age-related structural trajectories as well.

6.5.3. Towards a research hypothesis

When middle-aged and older adults are asked of their concerns about living to age 75, their greatest worries are not somatic. What mostly concerns people about growing old is memory function loss (Cutler & Whitelaw, 2005). If sensorimotor training leads to measurable changes in the adult brain, what about cognitive domains such as memory?

In 2006, it had been shown that a group of medical students studying for their exam exhibited transient grey matter increases in hippocampus, pointing to cognitive exercise-related plasticity at a level detectable by MR-morphometry (Draganski et al.). Still no longitudinal reports existed on the effects of memory training on brain structure in adulthood. Simultaneously, a large body of research had documented benefits of memory training in adult populations on the level of cognition (Rebok, Carlson, & Langbaum, 2007).

6.5.4. Cognitive plasticity

Cognitive plasticity can be defined as the ability to improve cognitive performance following training (Jones et al., 2006). A consistent finding in the memory-training

literature is that older adults can improve task-specific memory performance following training (Verhaeghen, Marcoen, & Goossens, 1992). We saw earlier that part of age-related decline in episodic memory is thought to stem from poor strategy use during encoding. As reviewed by Jones (2006), the mnemonic technique Method of loci (MoL) is one of the most commonly used procedures to improve performance in memory-training research. The MoL is thought to improve memory performance by providing the subject a strategic framework to facilitate encoding and retrieval of verbal information (Bower, 1970).

At the starting point of this thesis, it was unknown whether improving cognitive performance through memory training would be accompanied by structural alterations in the brain—at a level detectable by MRI. Thus, for the present thesis we wanted to study the effects of memory training on 1) cognitive performance, shown to be improved in previous studies and 2) changes in MRI-indices of cerebral grey and white matter, at this time not studied before. Finally, as relationships between the MRI-measures of aging brain structure and cognition had been documented in cross-sectional studies, we sought to investigate whether links existed between potential changes in brain and cognition following training.

6.5.5. Subjective memory impairment—further development of research hypotheses

"Knowing is not enough; we must apply. Willing is not enough; we must do."

—Goethe

It is proposed that in 2050, only 2.5 Norwegian adults will be in working age for every one person above retirement age (Brunborg & Texmon, 2010). With age being the strongest risk factor for cognitive decline and Alzheimer´s disease, it is evident that research initiatives targeting age-related decline and the development of interventions for individuals experiencing memory problems are of critical societal importance in the years to come. In this vein, a particular group of interest is middle-aged and older adults experiencing memory problems, but who are without objective impairment.

In a recent study, 19 individuals who contacted a Memory clinic due to persistent experiences of memory problems, but who were free of objective impairment were assessed on memory performance and functional MRI-estimates of neuronal activity during retrieval (Erk et al., 2011). Despite similar performance on the memory task, the subjects

experiencing memory problems showed reduced activation in the hippocampal formation, compared with adults without memory problems. Over-activation, as compared with controls, was found in the dorsolateral prefrontal cortex. The paper concludes that:

“the present results further support the concept of subjective memory impairment as a compensated pre-MCI [Mild cognitive impairment] state of the clinical manifestation of Alzheimer´s disease.” (pg. 850)

Other studies have indicated that subjective memory impairment is associated with increased hippocampal atrophy (Saykin et al., 2006; Striepens et al., 2010), a hallmark feature of Alzheimer´s disease pathology. Although controversial (Iliffe & Pealing, 2010), some studies indicate that the mere subjective feeling of persistent memory problems debuting in middle or old age represent a risk factor for developing Alzheimer´s disease (Jessen et al., 2010; Jonker, Geerlings, & Schmand, 2000). Depressive symptoms are also related to subjective memory complaints, with uncertain causality (Reid & MacLullich, 2006). As reviewed by Byer and Yaffe (2011), depression has in it self been associated with more than twofold increase in dementia risk. Still, to get a diagnosis of subjective memory impairment at the memory clinics, the clinician first rules out major depressive disorder.

At present, subjective memory impairment represents a substantial proportion of the total outpatient population at Norwegian memory clinics. 27% of all initially examined outpatients in 2010 fell under this category (Brækhus, et al., 2011). Moreover, 49% of referrals were under the age of 65 years (Brækhus, et al., 2011), pointing to a young group of individuals possibly at increased risk of a disease associated with great personal disability and socio-economical burden (Burns & Iliffe, 2009). Today, individuals with subjective memory impairment are left without any treatment option, but often ask what they can do about their problems.

Studies have indicated positive effects of cognitive intervention for older adults with memory problems (see Buschert, Bokde, & Hampel, 2010, for a review). Subjective memory impairment represents an interesting group of study in this regard. At the stage of amnestic mild cognitive impairment (MCI) and beyond, brain degeneration is so advanced that interventions aiming at remediation is less conceivable. For otherwise healthy elderly with subjective memory complaints, on the other hand, training might beneficial—at least

temporarily. Knowledge about factors influencing the potential for memory improvement is thus vital.

7. Main objectives and hypotheses

The overall objective of the present thesis was to investigate effects and correlates of an eight-weeks memory-training program on brain and cognition in middle aged and older adults. The three included papers seek to illuminate to what extent brain and cognitive plasticity are related in healthy and preclinical adult populations.

Paper I

Improvements in memory performance following Method of loci (MoL) training have been documented (Verhaeghen, et al., 1992). We wanted to assess whether this finding could be replicated in a middle-aged and older Norwegian cohort. For this purpose we designed a memory-training program aimed at improving verbal memory by implementing MoL-training. We hypothesized that memory training improves task-specific memory performance. That is, after memory training, participants will perform better on the skill targeted by MoL-training—recall of the order of concrete words (in this thesis denoted verbal source memory).

The second objective was to study whether training was accompanied by regional changes in the brain. In paper I, we focused on a well-studied property of the cerebral grey matter, namely cortical thickness. We hypothesized that memory training is associated with regional increases in cortical thickness compared with controls.

From a theoretical standpoint, we hypothesized that merely providing a new strategy for improving memory is not enough to lead to structural changes in the brain, as the increase in performance by learning the strategy in it self might only reflect flexibility within existing neuronal systems. It has been proposed that in order for environmental demands (cognitive exercise) to lead to a structural change in the brain, a *prolonged* mismatch between current functional level and demand is needed (Lövdén, Bäckman, Lindenberger, Schaefer, & Schmiedek, 2010). We hypothesized that the memory-training program needs a sufficient "dose" of practice (exercise) beyond everyday demands in order to lead to any measurable change in brain structure. In the words of Lövdén and colleagues (2010), we designed the current training program to "create a maximum impetus for change", by requiring participants to exercise (rehearse and later serially recall) up to 60 words in series five days

a week for eight weeks, using gradually shorter time-limits as a way to adapt task difficulty. Certain word-lists were to be remembered on following days, as well. We chose eight weeks of training as other training regimens had shown effects of intervention in similar, and even shorter intervals (Ilg et al., 2008). We chose not to expand the interval to three months as the above-mentioned juggling-studies, as we hypothesized that this could lead to significant motivational dropouts due to the rigorous, time-consuming nature of the training.

Finally, we assessed relationships between changes in memory performance and cortical thickness change for each participant. We hypothesized a positive relationship between memory performance change and change in cortical thickness.

Paper II:

Cross-sectional studies suggest brain structure age deterioration not only in cerebral grey matter, but also in the underlying white substance. Post-mortem data indicates that the cerebral white matter is the structure of the brain that declines most with age. Most, if not all data at the time was based on cross-sectional designs. It had not been shown whether longitudinal white matter changes are detectable within the present test-retest interval (about ten weeks). Neither had it been tested whether changes in white matter structure are modifiable through memory training. We tested the following hypotheses:

1. Longitudinal white matter changes are detectable by DTI as MD increases, FA decreases, or both—in line with the established age trajectories.

2. Training impacts changes in the DTI-indices, contrasting the age trajectories.

3. Greater memory training improvements are related to white matter changes as more positive FA increases, lesser MD increases, or both.

Paper III:

The second main objective of this thesis was to study whether the training program is of benefit for older adults experiencing memory problems, without cognitive impairment. Paper III evaluated the feasibility of the training program described in papers I-II for patients with subjective memory impairment referred to Oslo area memory clinics. We used the same program structure as for healthy adults, with a focus shift towards remediation and

support: Providing individuals experiencing memory problems tools and knowledge for improving episodic memory and at the same time an arena to meet other with similar challenges.

When reviewing the literature, I found little evidence regarding whom among this heterogeneous group of individuals would benefit from such a program. At the same time, findings from cognitive neuroscience pointed to structural MR-imaging as a tool for predicting performance improvements. For instance, volume decline in MTL structures are shown to predict 2-year reductions in episodic memory abilities in healthy elderly (Murphy, et al., 2010). Smaller left hippocampal baseline volumes correlated with progressing 5-year memory decline on the CVLT long-delay recall in another study (Tupler et al., 2007). Adjacent research fields further indicated that regional brain volumes correlate with relevant training benefits. In one study, Erickson and colleagues (2010) found that variability in greater pre-training volumes of the striatum explained 23% of the variance in performance improvement in a game version emphasizing cognitive flexibility.

These findings prompted the final objective of paper III: To study relevant pre-training brain volumes in the prediction of memory-training benefit in subjective memory impairment. In this study we used verbal recall performance from original and alternate versions of the CVLT-II as outcome measure of memory changes. The hippocampal formation was chosen as a brain predictor region of interest due to its well-known importance for episodic memory.

We hypothesized 1) that larger hippocampal size before training is positively related to memory changes following intervention, and 2) that possible relationships are specific to sub-regions within the hippocampus.

For this purpose we used a novel MR volumetry protocol enabling sub-region delineation of the hippocampus (Van Leemput et al., 2009). We performed separate analyses for seven hippocampal subfield volumes calculated by the MRI-analysis suite (FreeSurfer; see Methods) without specific a priori hypotheses regarding which particular region would be related. As memory complaints in the elderly are associated with increased depressive symptom load, we also assessed the effect of baseline depressive symptoms on training benefit. We tested for this by including a validated symptom scale as a covariate in multiple linear regression analyses.

8. Research questions

Paper I

1. Is systematic memory training associated with short-term changes in memory performance?

2. Is the training accompanied by regional cortical thickness changes compared with controls?

3. If so, are the changes in thickness and memory performance related?

Paper II

4. Are longitudinal changes in cerebral white matter detectable over a ten-week scan interval as MD-increases, or FA decreases, or both?

5. If so, does memory training impact these changes, and are they related to the memory performance outcome targeted by training?

Paper III

6. Is the memory-training program feasible for adults referred to a memory clinic with subjective memory impairment?

7. Are hippocampal pre-training volumes related to memory performance outcomes following training, and if so, are the effects selectively related to hippocampal subfields?

8. Do depressive symptoms at baseline co-vary with training benefit, and if so in which direction?

9. Methods

9.1. Design

For papers I-II, we employed a randomized controlled design in order to assess longitudinal group-differences in MR-imaging estimates and memory function. A particular advantage of the randomization procedure and control condition is that they reduce systematic bias and allow for questions of causality. Paper III also included longitudinal assessment of episodic memory performance, but did not include a control group. Thus, this paper does not address questions related to training-related brain plasticity or training efficacy.

9.2. Recruitment process

Recruitment and data acquisition for this thesis lasted from October 2007 to May 2008. In Papers I-II we recruited healthy individuals without memory complaints through a newspaper ad. In Paper III, we recruited outpatients with memory complaints at two major Oslo area memory clinics. All participants were community-dwelling middle-aged and older adults (ages 42 – 77) from the Oslo area. We started the outpatient recruitment first, and healthy volunteers were chosen to match the outpatient sample with regards to age, sex, and education. Figure 1 shows the complete recruitment flow for papers I-III.

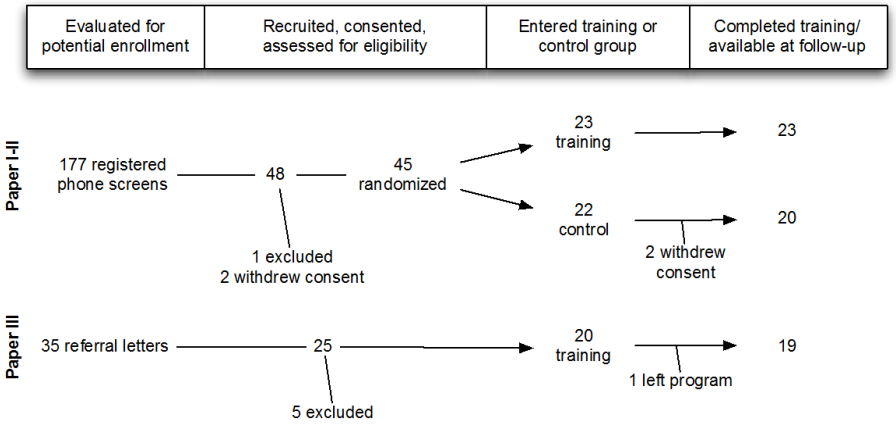

Figure 1. Recruitment flow. In papers I-II one participant was excluded due to poor neuropsychological profile. In paper III, five patients were judged ineligible due to probable dementia (three), post-traumatic stress disorder (one), and epilepsy (one).

9.2.1. Healthy volunteers

We announced the trial, and the most important criteria for eligibility in a single newspaper ad in a local Oslo newspaper (Aftenposten Aften). Inclusion and exclusion criteria are given in Paper I. Volunteers contacted us via e-mail or telephone. 177 interested subjects were screened with a structured telephone interview and registered in a database. 151 were deemed likely eligible based on the interview. The others reported obvious contraindications, such as too young or old, reporting to take medications known to affects CNS-function, such as antidepressant and anxiolytics, suffering from previous stroke or diabetes. Among those deemed eligible based on telephone screening, 46 subjects comprising 13 pairs matched on age, sex and education were initially invited to participate in neuropsychological testing and MRI after signing informed consent. One subject was excluded due to poor verbal memory performance (See Paper I). To replace this subject, we invited two other matching subjects to participate. The first withdrew consent after experiencing claustrophobia in the MRI scanner. The second subject withdrew consent after neuropsychological testing without providing any reason. Due to project time constraints, we ended the recruitment at this time. Thus, 45 participants were randomized, using a random-number generator in Excel, as follows: Each participant in each pair was assigned a random number. The participant with the higher number (between 0-1) was assigned to the intervention group. Two of the investigators performed the randomization together (Andreas Engvig and Torgeir Moberget). From the description it can be seen that the allocation of participants was not based on simple randomization, but followed matching of participants into pairs matched closely for age, sex and education. Due to the small sample size, this was deemed essential in order to obtain evenness of these variables across groups. The rationale for this is that these variables co-vary with performance and structural brain characteristics, particularly age. Significant group-differences on these variables would have implied brain and/or behavioral differences interfering with the interpretation of the main analyses of paper I and II.

As can be seen, papers I-II focused on the same set of participants. In both papers one participant had to be excluded due to inadequate MR scan quality. Furthermore, in paper II we had to exclude an additional participant from the analyses due to an incomplete diffusion imaging data set.

9.2.2. Memory clinic outpatient sample

The participant flow is presented in Figure 1. Subjective memory impairment comprises a heterogeneous group of individuals, and as the major goal of this study was to evaluate intervention feasibility and benefit, we applied rather broad criteria of inclusion: Subjectively experienced memory problems lasting less than 10 years, age [40, 80] years, MMSE > 25, estimated IQ > 85, no contraindications for MR-imaging, and evaluated as non-demented by the examining physician. Exclusion criteria: 1) No signs of stroke or other significant neuropathological lesions evaluated by a specialist in neuroradiology. 2) No major neurological or psychiatric disease.

We evaluated 35 referral letters for potential enrollment. 25 wrote informed consent after thorough verbal and written information about the project and were screened and tested on selected neuropsychological tests. 5 of the participants did not fulfill inclusion criteria (one had a history of post-traumatic stress disorder; one had epilepsy; three participants were evaluated as probable dementia by the examining physician). Thus, 20 eligible participants were invited to participate in memory training after completing MRI investigations. One participant left the program in the second week of the program. 19 completed the program and were included in the analyses.

9.3. Memory training intervention

The cognitive training literature is increasing, and several types of interventional approaches are available. Training programs to enhance memory in older adults can be divided into traditional and novel approaches (Rebok, et al., 2007). The latter denote training of ability. Ability training is herein defined as training that targets cognitive abilities or processes thought to *underlie* successful cognitive functioning: Examples of ability training include processing speed (Ball, Edwards, & Ross, 2007), working memory (Klingberg et al., 2005), and perceptual processing training—either auditory or visual (Smith et al., 2009). Ability training relies on computer software and is performed either in a group or an individual setting. Traditional approaches do not rely on use of sophisticated technology or computers, but instead focuses on teaching and working with strategies to enhance memory, and also teaching relevant information of importance for optimal memory and aging. In these interventions individual training is provided through “pencil and paper” assignments. Both approaches have strengths and limitations, which will be discussed. We chose to design a

group-based intervention based on traditional approaches for the present thesis. The program designed for this thesis is inspired by earlier work published by memory training pioneers such as Neely and Bäckman (Neely & Backman, 1993), and by a program designed for mildly cognitively impaired patients (Belleville et al., 2006).

Most of these programs have included some sort of memory encoding strategy to facilitate recall. In the present program we chose to put a particular focus on teaching the mnemonic technique known as the Method of Loci (MoL). The origin of the method goes back to ancient times and the poet Simonides of Ceos. (Cicero; Yates, 1966). Maybe the first componential analysis of the method and a description of its application in scientific terms were published in 1970 by cognitive psychologist Gordon Bower (Bower, 1970). As described by Bower, the method involves learning a series of landmarks (e.g., rooms in one´s home, places along the way to work) in a specified order. When the loci have been acquired, words or other information are linked to the various landmarks at the time of encoding. At subsequent retrieval, the landmarks are mentally revisited and the words linked to each locus are retrieved in serial order.

We chose to implement the MoL training as the main feature of the present intervention for several reasons:

1. It had an established theoretical foundation (Rebok, et al., 2007)

2. Multiple studies have shown that MoL training can improve task-specific memory performance in older adults (Verhaeghen, et al., 1992).

3. It is a well-studied approach for improving verbal memory in earlier training research, facilitating the comparison of results (Rebok et al., 2007).

4. To test whether memory training impacts structural brain changes, we reasoned that participants must have a certain dose of exercise that is continuously challenging, i.e., adapting to individual level and progress. Assignments increased in difficulty, including more words to be remember and reducing time limits throughout the program. MoL training enabled these hypothesized necessities through a low-cost approach. Note that for the outpatient sample we explicitly lowered the expectation on time limits, to avoid the training becoming a stressor.

5. MoL training did not require us to purchase and translate currently costly English software in order to test our primary objectives. Such requirements could have delayed, or even failed the undertaking of this thesis.

6. Of particular importance for paper III, MoL targets episodic memory function, a domain related to subjective concern in later parts of the lifespan (Cutler & Whitelaw, 2005).

Age-related reductions in successful MoL use have been reported in older samples. Singer and colleagues showed that individuals above 80 show less benefit from training in the MoL than do adults in their 60s and 70s (Singer, Lindenberger, & Baltes, 2003). We restricted the upper age-limit during recruitment to 80 years since we included a subclinical patient group where application of the method could be problematic.

The main aim of the program was to improve verbal recall by implementing MoL (Bower, 1970). In addition to the effects of learning a mnemonic technique, is should be noted that the program might have provided other context-sensitive active components. For instance 1) an arena for social stimulation and support (non-cognitive, possibly anxiety-lowering component), 2) teaching of practical memory principles (motivational component; multiple strategies), 3) targeting attentional function through introducing a technique for focusing attention; and by augmenting the awareness of the attentional demands in remembering.

Neely and Bäckman used the term multifactorial for training programs that also target attention and relaxation skills, in addition to mere encoding strategies (Neely & Backman, 1993). They reported that a multifactorial approach to MoL training resulted in larger effects of training in terms of both magnitude and maintenance compared with merely providing participants with the MoL technique (1993).

9.3.1. Program structure

The basic program structure and homework assignments were the same for papers I-III. The training program lasted for eight weeks, where each week consisted of a 1-h group session and 4 days of individual homework assignments. Group size ranged from 3 to 8 persons. The group sessions were held in a seminar room at the University of Oslo, and were lead by an instructor. There were two instructors involved in the present study (Andreas Engvig, and Vivi Agnete Larsen, professional students in medicine and psychology, respectively).

Each session followed a basic structure: Review of homework and positive feedback; focus on weekly and overall course goals; presentation of new didactic information; individual in-class memory training; and homework assignments. Power point summaries of the in-class curriculum were handed out for each session. The first classroom session lasted about 90 minutes; the following weekly sessions lasted about 1 hour. In the second to eight week the first 5-10 minutes were dedicated for the participants to discuss the previous weeks´ assignments, and to share and troubleshoot problems with the instructor and with each other. This discussion was lead and moderated by the instructor. Next, new didactic information was presented. After the presentation of the new material, for instance on remembering names and faces, participants were allowed to share concerns or insights regarding the current topic. This part of the session lasted about 15-20 minutes. During this stage effort was made to moderate the discussion so that all participants were allowed to participate. The last 20-30 minutes were devoted to in-class individual memory training using MoL.

In the appendix I describe how we taught the MoL and the additional content of each classroom session. Please see Table 1, Supplemental document 1 of paper III for an overview of the memory training program curriculum.

9.3.2. Individual MoL training

We made 40 exercises for individual MoL training, eight of which were performed in-class. The exercises are described in more detail in paper I, and example assignments are provided in the supplemental material of the same paper.

Briefly, for in-class training, eight standardized exercises were made. Each exercise included word lists of 10, 20, and finally 30 words to be remembered in serial order. Words were concrete nouns drawn for the Oslo Corpus of Tagged Norwegian Texts database (OC), provided by the Text Laboratory at the Institute of Linguistics, University of Oslo (http://www.tekstlab.uio.no/norsk/bokmaal/english.html). The word-lists did not include any of the words used in the experimental source paradigm (see below).

At the end of each classroom session, we gave participants four homework assignments to be completed on four of the six following days. The home exercises consisted of lists of concrete words to be remembered, and mostly followed the same structure: The participant

was told specifically to use the MoL to encode the presented word list in serial order. Next, the participant was told to read a short text (interference text) and finally serially recall the words to-be remembered. We gradually increased the home exercise difficulty by increasing number of words to be remembered, decreasing the total available time for encoding, and by using less visual words towards the end of the program. See supplemental material of paper I for an example.

9.4. Cognitive outcome measures

9.4.1. Verbal source memory

A hallmark feature of the MoL is the sequential organization of information to be remembered. After successful encoding, the subject can not only recall a list of words; but also recall the sequence of the words as they are encoded in series along the loci-route. At the time of designing this project, we were aware of no standardized test of measuring retrieval of the order in which previously word were presented. Thus, we developed an experimental memory paradigm to assess the effect of MoL training for the first two papers. The test design is described in paper I. Briefly, participants first had to make forced yes/no judgments whether a presented word had been displayed earlier (denoted verbal recognition). If a word was judged as old, the participants were prompted to indicate whether it was among the first, middle or last 5 words on the word list (denoted source memory). We hypothesized that Verbal recognition memory would not be improved by MoL training. However, for any correctly recognized word, we hypothesized that the MoL training group would improve on the ability to recall the source (order) of the word. Thus, we chose the ratio between correct source memory and recognition memory judgments as the main outcome variable target by MoL training.

Number of false alarms (recognize new word as old) was reported as group means and standard deviations in papers I-II. However, we did to correct the main outcome variable ratio for this. By correcting recognition hits in the denominator, a high level of false alarms might paradoxically (and artificially) have inflated the ratio by lowering the denominator. We considered the source/recognition ratio score to be fairly robust against response-style bias, as a larger number of recognized items should not affect the subsequent proportion of correct source judgments.

As can be seen above, we used the term source memory for recall of word order. The term source memory was used here since the subject needs to remember information about the context or source—in this case about the order—of previously presented stimuli. Surely there are other ways to operationalize and measure memory for source or context.

The reader might ask why the patient group in paper III did not perform this paradigm. After initial pilot runs on young adults, we implemented it for the healthy volunteer sample in papers I-II. The paradigm turned out to be feasible for all tested participants. The study reported in paper III was conducted simultaneously as the study on healthy volunteers, but these patients were not tested before training. I acknowledge that the current project could benefit from including the source memory paradigm described in papers I-II for the patient group in paper III as well.

9.4.2. California verbal learning test-second edition

In paper III, we assessed memory performance using a validated clinical instrument. Specifically, we employed the California verbal learning test-second edition (CVLT-II; Delis, Kramer, Kaplan, & Ober, 2000a). CVLT-II is the routine assessment form for verbal memory at the Neuropsychological unit at Oslo university hospital, Ullevål. The standardized neuropsychological test assesses various aspects of verbal learning and memory performance, and is described in detail in Paper III. In this paper, we used the long-delay (20-minutes) recall from the CVLT-II as an outcome variable. This metric has clinical relevance in the assessment of cognitive impairment and has been linked to hippocampal volumes in previous trials. This made the metric suitable for evaluating the usefulness of hippocampal volume as a MR predictor of recall change. Note that CVLT-II does not require the subject to remember words in order, but merely asks the subjects to recall all remembered words in any order. However, several aspects of the training program target the characteristics of this test. For instance, the program taught participants about various strategies to improve verbal memory, including concepts such as organization, association of related words, and visualization of information to be remembered. Most patients reported to use some of these strategies when performing the test.

9.5. Magnetic resonance imaging

Magnetic resonance imaging (MRI) is a non-invasive and non-ionizing technique used for internal organ visualization and measurement of tissue properties in vivo. Magnetic resonance image formation is complex, and introductory texts are available (Hashemi, Bradley, & Lisanti, 2010; Weishaupt, Köchli, & Marinček, 2006). The basis for MRI is nuclear magnetic resonance, a phenomenon in which charged atomic nuclei in a magnetic field absorb and emit electromagnetic energy. The electromagnetic energy is absorbed from radiofrequency pulses sent by an MR machine. These pulses have to be at the same wavelength as the nuclei of interest, hence the name resonance. The re-release of energy by the nuclei as they return to their initial energy state is detectable as oscillating currents—MR signals—by the MRI machine. MR signals are transformed into images and biological meaningful quantities by a computer. Additional gradient pulses are used in order to define the signals´ positions in three-dimensional space. For most imaging purposes, MRI relies on hydrogen nuclei (protons), present in the body as water, and inorganic molecules such as fat. Since, e.g., water content (and thus proton density) differs between tissues, the strength of MR signals varies depending on the tissue type (e.g., bone versus soft tissue). The time the protons use to return to their steady or equilibrium states also varies within tissues. Differences in these relaxation times are, put simply, the basis for differences in signal intensity—contrast—within soft tissues such as the brain. Signal intensity differences between grey and white matter are for instance used for calculation of cerebral cortical thickness, the brain outcome variable in paper I.

9.5.1. MR protocol and scanner

All imaging data were acquired using a 12-channel head coil on the same 1.5-T Siemens Avanto scanner (Siemens Medical Solutions, Erlangen, Germany) at Oslo university hospital, Rikshospitalet. The pulse sequences used for morphometric analyses were two repeated 3D T1-weighted Magnetization Prepared Rapid Gradient Echo (MP-RAGE) sequences. For diffusion imaging we used two repeated single-shot twice-refocused spin echo echo planar imaging pulse sequences. Each sequence included 30 independent diffusion sensitized gradient directions for measurement of diffusion in three-dimensional space.

Sequence details are found in the methods sections of the articles. Additionally, a T2-weighted fluid-attenuated inversion recovery (FLAIR) sequence and a T2 space sequence

were used to aid neurological examination. All baseline scans were subject to neuroradiological evaluation by an experienced specialist in neuroradiology. No significant lesions were reported, such as signs of stroke. However, we did not quantify or account for mild white matter lesions, which have been reported even in healthy forty year olds (Wen, Sachdev, Li, Chen, & Anstey, 2009).

MRI datasets for each participant were de-identified at the university hospital, and then transferred to Linux workstations at the Neuroimaging analysis lab at the Center for the study of human cognition, University of Oslo. Subsequent MRI-processing and analyses were performed at this lab with additional computing resources provided by the Titan high performance computing facilities at the University of Oslo (http://hpc.uio.no/index.php/Titan).

9.5.2 Cortical thickness and volumetry

In paper I and III we report data obtained from T1-weighted MP-RAGE pulse sequences. Partly due to differences in T1 relaxation times, white matter appears brighter (higher intensity) than grey matter, making separation of the two tissue classes feasible. The employed protocol was specifically developed to increase the contrast between gray and white (Brant-Zawadzki, Gillan, & Nitz, 1992), and is the basic acquisition protocol for brain morphometry studies. We ran two consecutive sequences, each lasting 6 minutes and 3 seconds, for each participant at each scanning session. Slices from each acquisition are compiled to three-dimensional volumes and subsequently averaged to improve signal to noise ratio. These average volumes were used as input volumes for the tissue segmentation procedures described in papers I and III.

We used the MR image analysis suite, FreeSurfer for automatic measurement of cortical thickness and regional brain volumes in the present project. FreeSurfer is documented and freely available for download (http://surfer.nmr.mgh.harvard.edu/).

8.5.2.1. Estimation of cortical thickness

For cortical thickness estimation, pial and white matter surfaces are created by FreeSurfer using spatial intensity gradients across the brain tissues and are therefore not simply reliant on absolute signal intensity. Therefore, the surfaces are not restricted to the voxel resolution

of the original data enabling inference of thickness differences at the sub-millimeter level (Fischl & Dale, 2000). The surface reconstruction and segmentation procedures are run automatically, but require supervision of the accuracy of the spatial registration and tissue segmentations. We performed no manual intervention to these procedures in the present thesis. This was avoided in order to reduce systematic bias due to human intervention. The accuracy of the procedure used have been validated against histological post-mortem analysis (Rosas et al., 2002), and manual MRI measurements (Kuperberg et al., 2003).

Furthermore, in a reliability study by Han and colleagues it was shown that user interventions are not necessary to obtain reliable thickness estimation, even across field strength, scanner upgrade and manufacturer (Han, et al., 2006). For acquisitions from the same scanner, thickness accuracy and test-retest reliability is particularly good. The reliability study by Han and colleagues included the same scanner type used in the present thesis (Siemens Avanto 1.5 T). Still, for one of the participants in paper I, FreeSurfer failed to fully remove the skull and dura, and the grey/white matter boundaries were obviously incorrect across multiple regions of the brain. This participant was excluded.

8.5.2.2. Estimation of hippocampal volumes

For paper III, we chose the hippocampal formation as a region of interest. For this purpose we performed volumetric segmentation of the whole hippocampal formation, and select hippocampal subfields, again using FreeSurfer. Accuracy of the whole hippocampal formation segmentation has been validated against manual MR-measurements (Fischl et al., 2002). Tae and colleagues (2008) found that FreeSurfer-estimated hippocampal volume correlated .85 with manual delineation. A recent reliability study suggested high test-retest volume overlap for the whole hippocampal segmentation with a dice coefficient of 0.87 ±0.02 in sample of older adults scanned on the same Siemens scanner (Jovicich et al., 2009). Test-retest reproducibility of left and right hippocampal volumes was 3.61% and 3.44%, respectively (Jovicich, et al.). Accuracy of the automated subfield procedure vas recently validated (Van Leemput, et al., 2009) and computes seven regional volumes: CA1, CA2/3, CA4/DG, presubiculum, subiculum, fimbriae, and hippocampal fissure, all included in analyses of paper III. CA denotes a cornu ammonis sector. DG denotes the dentate gyrus. Average dice coefficients are high (0.74) for the larger subfields, CA2/3 and subiculum,

whereas the two smallest structures, fimbriae and hippocampal fissure only reached an average dice coefficient of 0.52 in the validation study (Van Leemput, et al., 2009).

8.5.2.3. Longitudinal processing of cortical thickness

The present thesis includes longitudinal analysis of MRI scans. Longitudinal image processing is associated with certain methodological caveats, which needs to be accounted for. Several causes of bias in longitudinal processing have been identified, and are discussed in detail elsewhere (Reuter & Fischl, 2011). The first type of bias can be introduced during MRI acquisition; the most important issues being related to the use of different scanners, different software versions, and different scanning parameters. In the present study we sought to reduce these potential biases to a minimum, using the same scanner, scanner software, and scanning parameters for all participants and time points. Second, the researcher need to account for variations in imaging geometry over time due to varying head placement in the scanner. In the present thesis, we aligned acquisitions automatically using a standardized anatomical atlas online to ensure identical positioning of MR scan sections at the two time points (Benner et al., 2006; van der Kouwe et al., 2005). Next, we corrected the scans for gradient non-linearity distortions to further minimize image variation. For this purpose we used freely available and validated software (http://www.nitrc.org/projects/grad_unwarp; Jovicich et al., 2006). Bias can also be introduced during image processing when trying to obtain spatial correspondence of individual neuroanatomy across time points (Thomas et al., 2009). Normally, researchers try to achieve cross-time spatial correspondence by implementing various registration algorithms. Here, biased can be introduced when, e.g., the baseline (tp1) scan is used as a reference frame and therefore treated differently from the other time points. Ideally, all time points should be treated identically, such that image interpolation remains the same for each image. The FreeSurfer software used in paper I, includes a longitudinal processing scheme that seeks to solve this problem. For the case of two time-points, the first and second scan is registered to an average of the two. This is done to ensure unbiased analysis with regards to any of the time points. The scheme has been shown to significantly increase thickness measurement reliability across time-points (Han, et al., 2006). For paper I, FreeSurfer version 4.4 was used. Later improvements to the longitudinal processing stream has been added to later versions, mainly related to analysis of more than two time-points and by enabling manual intervention and editing of data.

9.5.3 Diffusion tensor imaging

The movement of water molecules within a tissue can be estimated by manipulating the strength, direction and number of radiofrequency pulses emitted by the MR machine. This basic feature is exploited in diffusion MR imaging, the principal method used in paper II. For diffusion tensor imaging (DTI), we fit a mathematical model to diffusion MR images – the diffusion tensor model. Tensor-related measures enable indirect inferences about white matter microstructure as physiological and pathological properties (e.g., axon-diameter, myelin-content or ischemia) of the tissue affect water movement differentially along the computed axes of the ellipsoid tensor. Water diffusion is, for instance, assumed more anisotropic along the fiber direction of major white matter tracts, enabling mapping of structural connectivity *in vivo*.

Two related phenomena seem to appear consistent in DTI-studies of aging (Chanraud, et al., 2010): The overall water diffusion within voxels of the white matter tissue increase with age, while the directional coherence of water displacement in white matter tracts decrease. Mean diffusivity is the most common metric to measure overall intra-voxel water diffusion, and is thought to reflect tissue barrier sparseness irrespective of direction. Directional coherence is reflected by the fractional anisotropy index (FA), which is a ratio from 0 to 1. Finally, by decomposing the diffusion tensor, additional metrics can be calculated as well, e.g., the average of the tensor´s smallest eigenvalues—the radial diffusivity (RD).

Excellent reviews on the basis for diffusion tensor imaging in the brain have been published (see, e.g., Alexander, Lee, Lazar, & Field, 2007). In the following I will discuss the issue of longitudinal processing of DTI-data. For the present study we used software utilities from the FMRIB software library (FSL; www.fmrib.ox.ac.uk/fsl) developed at Oxford by Steven Smith and colleagues (Smith et al., 2004). Specifically, we employed the tract-based spatial statistics (TBSS) protocol developed to ensure reliable cross-subjects alignment of diffusion tensor imaging data (Smith et al., 2006). The protocol enables multi-subject voxel-wise analysis of white matter microstructure by solving previous registration issues using, in part, an average FA skeleton derived from all subjects in a given study. Solid documentation exists for cross-sectional DTI-analysis using TBSS (Smith et al., 2007). However, a formal longitudinal TBSS-protocol is lacking (S. M. Smith, personal communication). A number of longitudinal DTI-studies report using TBSS. To the best of my knowledge, only a few investigators have included a longitudinal protocol description (see supplemental methods

in Scholz, et al., 2009). Inadequate spatial alignment of longitudinal imaging data has been showed to introduce bias affecting data interpretability (Thomas, et al., 2009). To solve the issue of across time-point registration, we wrote a longitudinal TBSS protocol combining tools available in the FSL toolbox. The primary aim of this protocol was to ensure unbiased alignment of time points. The protocol is described in sufficient detail in paper II to allow replication of the study. We are currently evaluating the protocol described in paper II against other procedures for longitudinal DTI-processing, and other labs are testing the present protocol on independent datasets as well (e.g., M. Strenziok, Arch lab, personal communication). We hope joint research efforts will lead to an established standardized protocol for longitudinal DTI-assessment in the near future.

9.6. Statistical analysis

9.6.1. Verbal source memory performance

We reported the results of longitudinal changes and group-differences in source memory performance in both paper I and II. The statistical analyses used were the same, and are described in the following. We performed repeated measures analysis of variance (ANOVA) to compare memory improvement between groups. Group-comparison of memory performance changes was performed to account for test-retest effects. We also performed independent-samples t-tests to compare baseline memory performance, and longitudinal change scores between groups.

9.6.2. Analysis of longitudinal changes in structural brain characteristics

9.6.2.1. Paper I: Global grey matter measures and regional cortical thickness

To reduce the probability of Type I errors, cortical thickness analyses were corrected for multiple comparisons. We used cluster size inference by means of Monte Carlo permutation testing implemented in FreeSurfer. In this approach the current data are tested against an empirical null distribution built from a given number of simulated datasets with an initial cluster forming threshold set by the user. In paper I, we chose 10.000 iterations and a cluster threshold of $p < 0.001$. In the present thesis, we were unsure what to expect in terms of effect size, as little longitudinal research existed concerning effects on training on cortical

thickness. We settled on using Monte Carlo permutation testing, but also performed exploratory analyses to evaluate uncorrected data. This included mapping all effects sizes across both hemispheres as p-value histograms, and using split-half validation of the training and control groups for evaluating consistency of uncorrected effects across subgroups.

9.6.2.2. Paper II: DTI-indices of regional white mater microstructure

A natural starting point to assess whether memory training impacts age-related change in DTI-indices of white matter would be similar to that of paper one: Investigate voxel-wise Group × Time interactions on MD and FA, testing whether mean longitudinal changes in the intervention group differed from that in the control group. However, as seen in paper II, the described voxel-wise analysis approach differed somewhat from the previous paper.

First, FSL provides different methods for cluster inference and multiple comparisons correction compared with FreeSurfer. Part of the statistical inference approach used by FSL and TBSS is denoted threshold-free cluster enhancement (TFCE). A common issue in most approaches to cluster inference is that the method used to control for multiple comparisons require the user to set an initial, often arbitrary cluster-forming threshold (e.g., $t > 4$; $p < 0.001$). The term "threshold-free" implies that FSL uses an algorithm that does not need to set any cluster-forming threshold (Smith & Nichols, 2009). In the case of TBSS, the statistical output is threshold-free cluster enhancement (TFCE)-corrected p-value maps; as compared with the correction performed in paper I which did involve setting an initial cluster-forming threshold. I acknowledge that the thesis would benefit from using the same statistical correction method in paper I and II.

Before we assessed specific effects of memory training on DTI-changes, we tested for baseline group-differences in FA and MD. While we failed to find group-differences in regional baseline thickness, whole brain DTI-analyses revealed regions of significant increased MD in the training group compared with controls. Thus, in the following interaction analyses we included baseline DTI-maps as voxel-wise regressors of-no-interest, an option provided in TBSS.

Then, in two separate analyses we tested Group × Time interaction effects on FA and MD across the whole white matter skeleton. This was done to model specific effects of training

using the same approach as the first paper—while accounting for baseline differences. We used TFCE for statistical inference after performing 10.000 permutations for relevant contrasts to identify clusters that are less likely to be a result of type 1 errors. The TFCE-corrected whole-brain p-value maps for MD and FA yielded no significant effects at the $p < 0.05$-level.

We now had several options. First, we could accept the null hypothesis at the expense of committing a type-II error. Second, we could go on with an exploratory investigation of uncorrected data from the whole-brain analyses at the expense of committing type I errors.

We chose a third option described under Research questions and in paper II: We evaluating smaller regions of the white matter skeleton to assess more localized, yet significant effects of training and decreasing the likelihood of type 1 errors. Instead of choosing an a priori region of interest, we first assessed whether any longitudinal changes in the diffusion data were evident across-groups. If short-term across-group changes would exist, we could then model effects of training within regions showing change—assessing whether training impacts the longitudinal changes.

Thus, we proceeded with voxel-wise analyses of MD and FA across-groups, testing whether the total sample showed brain regions of increased MD or decreased FA over time—in line with the age trajectory to be expected from previous studies. If an across-group change was identified in a smaller region of the brain, this would also reduce the number of statistical tests and the likelihood of Type I errors as compared with an exploratory approach on uncorrected whole-brain data.

As reported in paper II, longitudinal MD increases across-groups were evident in multiple regions of the brain. Reducing the corrected p-value threshold to 0.025 narrowed the result to anterior portions of the white matter skeleton with the bulk of voxels situated to the left of the midline. We used the anterior region as a mask to assess effects of training. This was supported by a recent publication suggesting that cognitive training affects age-related MD and FA-trajectories of anterior white matter (Lövdén et al., 2010). Again we assessed Group × Time interactions on FA and MD by cluster inference, accounting for baseline group-differences. The results are shown in paper II.

It is possible that the training affected other regions besides the identified cluster of the left anterior hemisphere. As reported above, the cluster inference method recommended by the

FSL developers failed to yield significant results at the whole-brain level. However, we re-ran the whole-brain analyses testing for Group × Time interactions on FA and MD using an alternative approach to correct for multiple comparisons which have been applied in studies of smaller sample sizes (see, e.g., Walhovd et al., 2010). Specifically, we used cluster-size thresholding on the statistical t-maps by clustering contiguous voxels by means of 26-neighbor connectivity in the white matter skeleton with a cluster-forming threshold of $t > 2.5$ (roughly corresponding to a p-value of 0.01 in this study). No significant results were found for MD. The figure below shows the group × time interaction analysis for FA yielding three significant clusters (corrected $p < 0.05$), suggesting increased FA in the training group compared with controls.

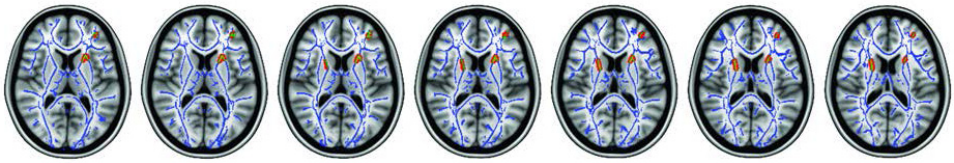

Figure 2. An alternative approach to cluster inference suggested significant FA increases in the training group compared with controls. Clusters of voxels (> 100) with significant ($t > 2.5$) FA increases in the training group compared with the controls are shown in red-green as they appear in the WM skeleton shown in blue. No significant clusters of voxels (> 100) were found with the inverse relation (more FA increase in controls compared with the training group). Radiological convention; the left side of the brain is to the right.

The largest cluster overlapped the originally reported cluster in the left hemisphere, intersecting the uncinate fascicle and anterior thalamic radiation (see paper II). The second cluster overlapped the anterior thalamic radiation in the contralateral hemisphere, mirroring the first cluster. A third cluster was found in the left frontal pole.

9.6.3. Relationships between brain changes and source memory change

We used the number of correct source judgments divided by number of correct recognition judgments as a ratio measure of verbal source memory performance targeted by training. This ratio is referred to as source memory in the following for brevity.

In paper I, we tested brain-cognition relationships by regressing cortical thickness change on source memory change (tp2 – tp1), for each subject, at each cortical region were an effect of training was identified. We also reported results for baseline adjusted memory

changes. Scatterplots were used to demonstrate individual changes in thickness and memory (See paper I).

In paper II, we assessed relationships between FA change in the region were we identified an effect of training and source memory changes. Follow-up analyses were performed on RD and AD. We chose to use spearman´s rho correlation coefficient analysis. I acknowledge the differences in analysis approach between paper I (linear regression which for two variables equals a bivariate Pearson´s correlation) and paper II (Spearman´s rho correlation). Briefly, rerunning the data in paper II using Pearson´s correlation coefficient analysis yielded comparable, significant results. For the training group (n = 21), Pearson´s estimated relationships between source memory change and FA and RD change were (FA: r = .48, p = .029) and (RD: r = -.46, p = .035), respectively. Please confer paper II for the results of the significant Spearman´s analyses.

9.6.4. Regression models of hippocampal volume and recall change

In the third paper, we tested relationships between pre-training hippocampal volumes and changes in verbal recall by means of multiple linear regressions. Shapiro-Wilks test´s were used to test the dependent and independent variables for normality. Left and right hippocampal volumes were included as independent variables in separate analyses. Follow-up analyses were performed for each of seven hippocampal subfield regions to assess subregional specificity. We chose change in long-delay verbal recall as dependent variable in all analyses. Note that the change score was statistically adjusted for baseline performance before being submitted to linear regressions. This was done to account for variability in starting level on memory performance change. Also, age, and sex were regressed out in all analyses. The primary objective was to assess regional brain volumes as predictors of recall change. The hippocampal formation was chosen due to its known importance in episodic memory. This reduced the number of statistical tests considerably compared to a whole-brain voxel-wise approach. We also tested whether depressive symptom load co-varied with recall change. Specifically, we included the Geriatric depression scale (GDS) as an additional independent variable in the regression models, and assessed the contribution of this variable to the prediction models.

9.7. Research ethics

The Regional ethical committee of Eastern Norway approved the current research project (REK project number 1.2007.911). Written informed consent was obtained from all participants included in the study prior to examination. As the project involved MR examinations including older adults, we ensured clinical neuroradiological evaluation of all participants through our collaboration with the Oslo university hospital, Rikshospitalet and specialist in neuroradiology and head of department, Paulina Due-Tønnessen. This was important as some of the participants were referred to a memory clinic with suspected cognitive disorder (paper III). None of the participants required follow-up due to MR-pathology. An ethical consideration relates to participation benefit of the no-contact control condition (paper I-II). Participants allocated to the intervention were offered a training program possibly associated with better cognitive performance and other potential social benefits, whereas the control group did not receive any intervention. As a compensation for this benefit mismatch we provided a two-week training and educational program for all controls after the end of the study phase. Also, all participants received an honorary for their time, and transportation costs.

The datasets of this project have been available and inspected by the core members of the research group. All analyses reported in the papers have been performed or supervised by more than one person. The co-authors are psychologists, physicians or professional students within those disciplines and thus legally obliged to professional confidentiality by the Health personnel act.

10. Main results

Paper I

The study was designed to assess effects of intensive two-months memory training on verbal source memory and cerebral cortical thickness. The numbering below refer to the research questions on page 31.

1. We first tested for effects of training on source memory performance. Repeated measures ANOVA, testing whether source memory change differed between groups, revealed a significant Group × Time interaction ($F(1, 39) = 14.8$, $p < 0.001$). Follow-up analyses revealed a significant improvement in the training group (paired t-test, $t = 7.04$, $p < 0.0001$), while for the control group a trend improvement on the same scores failed to reach statistical significance ($t = 1.73$, $p = 0.10$). Cohen´s *d* for source memory change comparing the training and control group was 1.2. This is an effect size of considerable strength (Cohen, 1988), and implies that the average improvement score of the training group is at 88th percentile of that of the control group.

2. Voxel-wise GLM, testing whether regional changes in thickness differed between groups, revealed increased cortical thickening in four regions in the training group, compared with controls: Effects were found in right anterior insular cortex, bilateral orbitofrontal cortices and in the fusiform gyrus. We used a split-half strategy for result validation; the results remained consistent across four sub-group analyses. Cohen´s d ranged from 1.3 (right insula) to fusiform 1.2 (right fusiform). The effect in the insular cortex also remained significant using the Monte Carlo permutation testing procedure.

3. Baseline-adjusted change in the source/recognition ratio correlated positively with thickness change in the right fusiform ($r = .39$) and lateral orbitofrontal ($r = .32$) cortices. Re-test thickness in the left lateral orbitofrontal cortex correlated with re-test source-memory performance ($r = .41$).

Paper II

The purpose of this paper was to assess effects of memory training on changes in white matter microstructure. The sample is the same as in paper I.

4. We detected longitudinal increases in mean diffusivity (MD) across groups (corrected $p < 0.05$), with the strongest effects in anterior white matter regions. Rate of MD change correlated with age (Spearman's $\rho = 0.38$, $p = 0.014$), suggesting larger MD increases with advancing age. No significant longitudinal changes in fractional anisotropy (FA) were found across groups.

5. Within frontal regions were MD increases were identified (corrected $p < 0.025$), we found a significant Group × Time interaction effect on FA, indicating a *relative* FA increase for the training group. The effect was driven by a significant FA decrease in the control group (paired t-test, $t = – 3.6$, $P < 0.002$). While a trend FA increase was found in the training group, the result failed to reach statistical significance ($t = 1.4$, $p = 0.17$). In the training group, greater FA change was related to better source memory improvement (Spearman´s $\rho = .496$, $p = 0.022$). Eigenvalue analyses showed that the relationship with source memory change was specific to change in radial diffusivity ($\rho = – .547$, $p = 0.010$), and not to axial diffusivity ($\rho = .026$, $p = 0.910$).

Paper III

The objective of the third paper was to assess whether the training program was feasible for memory clinic outpatients with subjective memory impairment. Knowledge about who might improve from training was scarce, and we evaluated whether hippocampal formation pre-training volumes could predict training benefit. We used a novel protocol enabling hippocampal subfield delineation (Van Leemput, et al., 2009) in order to study specificity of potential effects.

6. The memory-training program was feasible for a sample of 20 memory clinic outpatients with subjective memory impairment, including high participation (average 96.7%, all above 80%) and low-drop out rates (5%; 19 of 20 completed).

7. Pre-training volumes of left, but not the right, hippocampal formation correlated positively with change in CVLT-II long-delay recall ($\beta = .73$, $p = 0.019$), corrected for baseline performance. This indicates that greater size of the whole hippocampus before training was related to greater improvement in long-delay recall following training, independent of baseline performance. Subfield analyses suggested that the relationship with recall improvement was selectively related to the CA2/3 ($\beta = .66$, $p = 0.009$) and CA4/DG ($\beta = .67$, $p = 0.013$) volumes, independent of age, sex and baseline performance. We found no significant relationships for the other tested subfields and recall change.

8. Results from multiple linear regressions including both hippocampal volume and geriatric depression scale (GDS) showed that more depressive symptoms at baseline were associated with greater recall improvements following training ($\beta = .41$, $p = 0.047$), independent of hippocampal volume. There was no correlation between hippocampal size and depressive symptoms. We failed to assess GDS-score longitudinally, and cannot conclude whether training was associated with changes in depressive symptoms.

11. General discussion

The present thesis reports on a longitudinal neuroimaging study of memory training. Subjects were middle-aged and older adults with and without memory problems. We used state-of-the-art MR-imaging analysis in an attempt to link brain and cognitive plasticity in aging. New insights were learned. Papers I-II suggested that effects of memory training include not only cognitive improvements, but also structural changes in the human brain across MR-modalities. Structural changes correlated with participants' degree of change in the relevant cognitive measure, pointing to a link between the changes in brain and cognition.

Paper III suggest that specific hippocampal volumes and depressive symptoms pre-training independently predict memory-training benefit. The major findings of three papers are discussed. I address methodological limitations of the present studies and within the field. Pointers to future research are given throughout.

11.1. Memory training is accompanied by improvements in task-specific memory performance

In line with earlier research, Method of loci (MoL)-training in healthy middle-aged and older volunteers led to improvements in verbal source memory (Cohen´s d = 1.2). Verhaeghen and colleagues (1992) found a mean-weighted d of 0.80, including 12 studies of MoL-training in healthy older adults. The present study included on average a younger group of adults, a longer training program, and more exercise sessions than several of those analyzed by Verhaeghen. These features may all serve to strengthen the effects of the current training program.

A methodological issue regarding MoL-training is whether the participants in fact used the technique as instructed. Studies have shown that some older adults instructed in MoL use other spontaneous strategies instead (Verhaeghen & Marcoen, 1996). Most importantly, this relates to the validity of the conclusion that MoL-training lead to improvements in task-specific performance. To investigate this, we performed telephone interviews two to three weeks after the end of the study phase, asking whether participants had used the MoL as

instructed. There were four answer categories assessing to which extent the participants had used the MoL: Never, sometimes, most of the time, always. 86.4% and 81.8% answered that they had used the technique all or most of the time during the homework and individual in-class training, respectively. Thus, it seems, most, but not all used the technique as instructed. Surprisingly, one person reported to never have complied to use the technique, even though instructed to do so for eight weeks. The participant told that he had developed his own technique, creating orderly stories of the words to be remembered. The subject´s memory performance change score adjusted for baseline performance was 0.5 SDs below sample mean. The reported Group × Time interaction of paper I (pg. 5), remained significant after excluding the non-complying participant ($F(1,38) = 14.3$, $p = 0.001$).

MoL is possibly the most researched mnemonic strategy for improving memory function in older adults (Jones, et al., 2006), and performance effects have been replicated (ibid.). For the present thesis, we used MoL-training as a well-established model to investigate memory-training effects on brain structure, and how they might relate to task-specific improvements in cognitive function. Still, commercial computer training programs are currently being developed by others (Fernandez, 2010). The given rationale for these approaches is that traditional memory training, like the present, has mostly failed to report transfer effects to near and far cognitive and real-life functions (Brenes, 2003; Owen et al., 2010). The ACTIVE randomized controlled multi-center trial including 2832 participants showed that memory training teaching mnemonic strategies and providing pen-and paper assignments failed to transfer to other cognitive domains or instrumental activities of daily living (IADL; Willis et al., 2006). Although of similar length as the present program, the ACTIVE trial failed to include any homework assignments, which could be of importance for training gains. Yet—in the same study—computerized speed-of-processing training led to significant improvements in everyday abilities, including IADL-function and safer driving performance (Ball, et al., 2007), and reduced risk of depression at one- and five-year follow-up (Wolinsky et al., 2009).

The present studies included an assessment interval spanning nearly 10 weeks. We cannot conclude based on the present results whether the effects of MoL training on cognitive function, or brain structure, are maintained. Similar training program have shown sustained effects on memory performance at two-year follow-up (Ball et al., 2002), with some maintenance even five-years after intervention initiation (mean effect size 0.23; Willis, et

al., 2006). At present (January 2012), we are recruiting participants from paper I-II for follow-up assessments.

11.2. Memory training is accompanied by regional changes in brain structure, partially related to memory performance

In a recent opinion article on the future of cognitive intervention research, the authors recommend researchers to use multiple benchmark measures to determine training-effect and generalizability, including structural neuroimaging (Ranganath, Flegal, & Kelly, 2011). In the present studies we report effects of memory training as structural changes in cortical thickness and white matter integrity, in line with our hypotheses.

Permutation testing revealed significant cortical thickening in the anterior insular cortex in the training group compared with controls (paper I). Mean difference in thickness change between groups was 0.049 mm. Split-half validation further confirmed consistent cortical thickening in the right fusiform gyrus, and in the lateral orbitofrontal cortices compared with controls. The changes were partially related to increased memory performance at retest. Data from paper II further suggested that training impacts frontal white matter changes. Here, we found that more positive FA change in the training group predicted greater memory improvement. 40% of the voxels showing an effect of training encompassed the uncinate fasciculus association tract. This tract is known to project forwardly to the left orbitofrontal cortex where we found a training-effect on thickness (see fig 8. pg 10 of paper II).

Summarized, I believe the converging effects across MR-modalities, and the relationships with relevant performance strengthen the significance of the present findings. The present results implicate that plasticity of both brain and cognition extends beyond middle age. The brain-cognition change relationships are exciting, and point to a link between inter-individual variability at the two levels of enquiry.

Since we started the current research project, several research groups have documented structural brain changes in response to various training regimens (see Johansen-Berg, 2011, for a recent review). Most of these studies report grey matter increases following training, in line with the findings of paper I. Studies including follow-up assessments have, however, indicated that grey matter increases are in part temporary and decrease after an initial

learning or training phase (Draganski, et al., 2004; Draganski, et al., 2006; Taubert et al., 2010).

A weakness of the present research project is the lack of long-duration follow-up. Long-term assessment of participants is needed to illuminate the temporal extent and dynamics of memory training effects. An interesting future enquiry is to include more scanning sessions during training, which could allow dose-response interpretations. Also, studies show long-term effects of brief post-intervention booster-training on cognitive performance (Willis, et al., 2006); whether booster training further impacts the brain trajectories remains unknown.

The earliest MRI-studies of brain plasticity mostly included younger adults. More recently, a number of investigators have begun investigating brain plasticity in older adults (Erickson et al., 2011; Lövdén, Bodammer, et al., 2010; Lövdén, Schaefer, Noack, Bodammer, et al., 2011; Wenger et al., 2012). These papers point to a new wave of intervention research using longitudinal imaging—moving from mere detection of training effects to investigating age-limits of brain plasticity and the possibility of alleviating structural age-related decline.

A recent study by Wenger, Lövden, and colleagues (2012) represents the first intervention study of older adults using cortical thickness as an outcome variable following our paper from 2010. Here, both young and old adults underwent a four-month navigational training program. The program required participants to learn navigating in a virtual zoo. The young group exhibited cortical thickness increases in left precuneus and paracentral lobule following intervention, while the old exhibited no significant thickening in these areas. An old control group showed significant frontal lobe thinning in one area, whereas the training group showed no change. The authors write that their findings seem to contradict our study from 2010 (pg.3395); further, they argue based on their findings that normal aging reduces the potential for experience-related change. The two studies are, however, hardly comparable in terms of training-task and content. Also, and in line with a growing body of literature indicating a role of the human hippocampus in navigation-related plasticity (e.g. Maguire 2000; 2006), the same research group reported preservation of hippocampal volumes in the same training sample compared with controls in another article (Lövdén, Schaefer, Noack, Bodammer, et al., 2011). In the latter study the authors conclude in the abstract: “[…] sustained experiential demands on spatial ability protect hippocampal integrity against age-related decline.”

Regardless of the conflicting conclusions of navigation-related brain plasticity on cortical thickness and hippocampal volumes in older adults: Only a replication memory training study could delineate whether the present intervention (paper I) alter the structural trajectories of cortical thickness in a stable manner across different samples.

11.3. Hippocampal volumes predict training-benefit

In the third paper we investigated whether hippocampal volumes were related to performance outcomes following memory training in subjective memory impairment. Little was known about the correlates of individual training benefit in this heterogeneous group. In current study we used pre-training volumes of the hippocampal formation as a priori region of interest to study changes in verbal recall. The hippocampus is heavily involved in episodic memory (Squire, Stark, & Clark, 2004), and supported the selection of this brain region. We chose change in long-delay free recall performance on the CVLT-II as an outcome measure of training.

A pressing question is to which extent the CVLT-II change score reflects intervention benefit. For instance, the mean change in recall performance was 3 points. Although statistically significant, the *clinical* significance of this result remains unclear. To assess whether the recall improvements reported in paper III exceeded test-retest effects, I compared the CVLT-results with 19 age-, sex-, and education-matched healthy controls. The subjects are from the control group in papers I-II, which simply took the test at second time without training. I performed repeated measures ANOVA testing whether long-delay free recall change differed between the outpatient group and controls. The result revealed a significant Group × Time interaction ($F(1,36) = 6.249$, $p = 0.017$), suggesting greater recall increases in the outpatient group. This is an uplifting result, indicating that the recall increases are past test-retest effects. Still, if the present intervention is to be offered as a treatment option in the clinic, a randomized controlled study is essential to conclude on and delineate the possible effects of training, both on the short- and long-term.

Inference about brain and behavior is often performed through studying differences in central tendencies or averages (e.g., brain atrophy) in response to an intervention (e.g., disease modifying behavior). Another approach to linking brain and cognition is inter-individual variability. Paper III provides correlation evidence suggesting that inter-individual variability in hippocampal volumes predicts verbal recall changes following

memory training. It is still uncertain whether pre-intervention brain characteristics can be used clinically as a biomarker to assess individual potential for training and in selecting candidates for treatment regimens. Logistic regression with leave-one-out validation in a larger clinical cohort is a possible next step in this regard.

From a clinical viewpoint, I think the findings are first and foremost important in that they show potential of MRI to foresee treatment outcomes in a group of individuals were few, if any, predictors of outcome were currently available. At the memory clinic, MRI assessment is part of the routine workup (ref). In the present study, we employed a fast and automated volumetry procedure that uses rather standard T1-weigthed MR-images. A future line of research is to assess whether the present segmentation procedure could be implemented in a clinically meaningful format for assessment of outpatients with memory problems.

In a study including seven-year follow-up of older adults with and without subjective memory impairment, 15% of those *without* subjective impairment showed significant memory decline at follow-up (Reisberg, Shulman, Torossian, Leng, & Zhu, 2010). In contrast, 54.2% of those *with* baseline subjective memory impairment had declined in memory, and had a higher incidence of Alzheimer´s disease (AD). This supports the concept of subjective memory impairment as a meaningful entity in neurogeriatrics. With the possibility that subjective memory problems may indicate increased risk of AD, an important, but maybe not obvious limitation must be stressed: We have no reason to believe that cognitive training programs—such as the present—can prevent neurodegeneration and dementia. My position is instead that paper III and other work implicate that for older adults who experience memory problems, but who are otherwise healthy, cognitive training programs could temporarily improve cognitive performance, in part through alleviating affective constraints on memory.

11.4. Further considerations on brain plasticity research

11.4.1. Studies of brain plasticity in development and aging

Recently, Woollett and Maguire published a longitudinal investigation of navigational-related grey matter plasticity in London taxi drivers (2011). The results showed structural increases in the posterior hippocampi over a three-year period, supporting the findings from cross-sectional studies (Maguire, et al., 2000; Maguire, Woollett, & Spiers, 2006). The

authors acknowledge the longitudinal studies conducted in recent years, but point out methodological issues (pg. 2109): For instance, Woollett and Maguire interpret studies involving children and young adults as "potentially conflating brain development with learning". As an example, normal development literature including adolescents and young adults indicates hippocampal volume increases with advancing age (for a review, see Toga, Thompson, & Sowell, 2006; Østby et al., 2009). This observation follows the direction of the intriguing findings from non-randomized brain plasticity studies in young adulthood (e.g. Draganski, et al., 2006). Future studies of plasticity in childhood and adolescence should incorporate randomized controlled designs to reduce developmental confounders.

Studying remediation approaches in older age ranges could be beneficial in this regard, as the aging trajectories of several brain structures do not follow, but contrast the hypothesized effects of most interventions targeting age-related decline (see Erickson, et al., 2011; Lövdén, Schaefer, Noack, Bodammer, et al., 2011 for examples). In this vein, an unknown factor of causality needs further attention in plasticity research: Are the training-related effects on aging brain structure a result of "positive" plasticity, or are the training alleviating negative effects of age? In the present studies, we found group by time interaction effects on brain structure. The combination of training-related increases and control decreases reported here points to a mixture of two phenomena. Whether training-related structural adaptability occurs in concert with age-related alleviation through mutual or exclusive processes needs further investigation.

11.4.2. Conflicting results in studies of white matter plasticity

The sum of cross-sectional and longitudinal imaging studies on white matter plasticity supports the idea that practice and experience shape specific anatomic features of white matter microstructure (Jäncke, 2009). Nevertheless, and as pointed out by Jäncke (2009), some of the findings in this literature are contradictory. Again, the degree of inconsistency in findings depends on which part of the life span is being studied, but also the task or training-regime involved.

DTI-studies of experience-dependent changes in white matter have primarily investigated effects of either sensorimotor or cognitive training. The largest inconsistencies appear in studies of young adults, and most controversy regard studies on sensory-motor practice; cross-sectional designs in particular. Several studies of white matter plasticity have focused

on behavioral expertise (e.g. musicians or ballet dancers) in young adults, primarily in their 20s (Hanggi, Koeneke, Bezzola, & Jäncke, 2010; Imfeld, Oechslin, Meyer, Loenneker, & Jäncke, 2009) and early 30s (Bengtsson et al., 2005). In 20- and 30-year olds, properties of the white matter microstructure may not yet have reached their respective developmental plateaus. According to Westlye and colleagues (2010), mean age of global FA maxima was 29.1 years; for MD, mean age of global minima was 35.7 years. Resembling the critique on grey matter studies by Woollett and Maguire (2011), one of several explanations contributing to the discrepancies is as follows: In young adults, different developmental mechanisms unrelated to the intervention or behavior at study are at play. I acknowledge that other confounders are associated with studies of aging as well; but I merely point out that naturally occurring developmental differences could introduce bias in the interpretations of cross-sectional studies of experience-related structural adaptability.

The studies by Hänggi, and Imfeld (both report decreased FA in experts compared with controls), suggest an opposite relationships between DTI-measures and behavior compared with, for instance, the Bengtsson study (higher FA correlates with more practice among experts). In support of the first two studies is a recent longitudinal study suggesting reduced FA over multiple time-points in response to sensorimotor training in young adults (Taubert, et al., 2010). Another aspect—the diffusion tensor modeling itself—also needs to be taken into account: Overall reductions in FA (Pierpaoli et al., 2001; Wiegell, Larsson, & Wedeen, 2000) can be a result of increased coherence in an individual fiber population of anatomic regions containing intra-voxel fiber crossing (Tuch, Reese, Wiegell, & Wedeen, 2003). Future studies could take into account the localization of effects, and fit crossing fiber models (Jbabdi, Behrens, & Smith, 2010) to areas with known multiple fiber directions.

When looking to the longitudinal *cognitive training* studies, I am only aware of studies reporting FA increases in response to training—in all phases of the life span (Keller & Just, 2009; Lövdén, Bodammer, et al., 2010; Takeuchi et al., 2010). Meditation practice—as a proxy of mental exercise—was accompanied by FA increases in another longitudinal trial (Tang et al., 2010). In relation to cognitive aging in particular, these findings fit with studies reporting positive relationships between FA and cognitive performance (Kennedy & Raz, 2009; O'Sullivan et al., 2001). The findings support an intuitive, but simplified idea of "higher FA is better" in terms of cerebral aging and cognition. Surely, exceptions exist.

11.4.3. Neurobiological mechanisms

Common to papers I-II and other MRI-studies of brain plasticity in humans is the uncertain nature of the underlying cellular mechanisms (May & Gaser, 2006). Surgical resection due to epilepsy has provided some histological basis of MRI-correlates of human brain tissue (Concha, Livy, Beaulieu, Wheatley, & Gross, 2010), Further enquiry on similar relationships in the healthy is limited due to the obvious non-invasive requirements of human research. Nevertheless, several authors have speculated on possible neurobiological explanations underlying MR brain plasticity; mostly based on studies on rodents undergoing training or in so-called enriched environments (Draganski & May, 2008).

With regards to grey matter increases (volume, thickness), animal studies suggest that experience-related grey matter increases result from an array of morphological alterations. These include local synaptic events such as the formation of new connections by dendritic spine growth (Holtmaat, Wilbrecht, Knott, Welker, & Svoboda, 2006), and change in the strength of existing synapses (Chklovskii, 2004; Chklovskii, Mel, & Svoboda, 2004; Holtmaat & Svoboda, 2009). Creation of new spines is shown to continue into adulthood (Yasumatsu, Matsuzaki, Miyazaki, Noguchi, & Kasai, 2008), happens within days (Xu et al., 2009), and is detectable in the mature neocortex of rodents (Yang, Pan, & Gan, 2009).

As reviewed by (Draganski & May, 2008), other mechanisms related to training-related plasticity include changes in gene expression, and protein synthesis. In line with this, a recent study found that navigation-related hippocampal volumes increases in rodents were accompanied by increased staining for a neuronal growth-associated protein, but not for glio- or angio-genesis (Lerch et al., 2011).

Comparable to grey matter studies, DTI-indices of white matter do not allow specific inferences of the underlying biology (johansen-berg; Beaulieau). One putative mechanism to modulate age-related FA-decreases, however, is myelin breakdown (Bartzokis, 2004). In contrast, a role for experience-dependent CNS-myelination is supported by experimental studies suggesting that neuronal firing stimulates myelination in an activity-dependent manner (Demerens et al., 1996; Ishibashi et al., 2006; Wake, Lee, & Fields, 2011).

Two findings in paper II supports that myelination might be involved in use-dependent changes in FA: 1) In the area where we report a training-effect on FA (paper II, fig. 6), eigenvalue analyses suggested less radial diffusivity (RD)-increases in the training group

compared with controls. 2) RD change showed a stronger relationship with memory improvement compared with FA change. As RD changes have been linked to myelination (e.g., Song, et al., 2002), our findings match the view that the white matter plasticity is coupled to myelination-related processes (Fields, 2008).

11.4.4. How does the MRI-observed brain changes fit with our understanding of the plasticity?

Most longitudinal studies of experience-related structural MRI changes report relatively limited intervals (weeks or months). In one study, the authors reported regional increases in grey matter intensity (interpreted as mass) during 15 days of learning medical school curriculum (Ceccarelli et al., 2009). Others have reported similar intervals: e.g., four-week interval, 11 hours meditation practice (Tang, et al., 2010); two-week interval, 15 minutes daily mirror reading (Ilg, et al., 2008). The study by Ilg and colleagues (2008) comprised a total of 3.5 hours of mirror-reading; the authors found grey matter signal increases of 5% in the right dorsolateral occipital lobe. Recent evidence even indicates that structural changes detectable by MRI can be induced in a matter of hours of perceptual discrimination learning (Kwok et al., 2011).

Basic research confirms that neuro-structural adjustments happens within days in response to learning (Xu, et al., 2009). 5-days of maze training in mice led to increases of hippocampal volume by 3-4% in one study (Lerch, et al., 2011). The increases measured by MRI were in turn correlated with neuronal remodeling—revealed by post-mortem histology. An compelling study of tool learning in monkeys included ultra-high resolution MRI-data from 18 individual scanning sessions over a six week interval (Quallo et al., 2009). The study demonstrated grey matter increases that were evident after one week of training, and related to individual learning success. The direction and extent of results were comparable to the findings seen in human studies.

The findings of macro-scopical brain changes over relatively limited intervals challenge our understanding of plasticity, and point to a need for systematic model investigations. A reasonable position is that a supply-demand mismatch must be present to trigger plastic structural brain changes, and it can also be argued that this needs to be of some duration. To remain efficient, a neuronal system must also remain stable to maintain established knowledge and skills, and hence a theoretical position is also that there must be

considerable slowness in large-scale plastic responses. Arguably, too high degree of brain plasticity seems as dysfunctional as total brain inertia (Hertzog, et al., 2008); a neurocognitive system under permanent renovation could easily be maladaptive and too energy consuming (Lövdén, Bäckman, et al., 2010).

I believe that firm evidence holds that training-related brain changes are indeed plausible. Instead, I hereby argue that critical questions relate to what extent MRI can detect such changes in a stable and meaningful manner. Further, how do methodological limitations—such as biased longitudinal processing—restrict MR-imaging from advancing our knowledge about human brain plasticity?

11.4.5. Reducing bias in MRI-studies of brain plasticity

In the present thesis, converging findings across MR-modalities and correlations with behavior bring greater confidence to the present conclusions. Still, replication of studies on brain plasticity in adulthood is warranted. In a recent, comprehensive review (Johansen-Berg, 2011), only one of the identified studies came out negative (Thomas, et al., 2009).

Is it possible that even more contradictory results on the effects of training on brain plasticity exist, but have failed to reach the light of the scientific community due to the file-drawer effect (Scargle, 2000)? The paper by Thomas and coworkers (Thomas, et al., 2009) points to an important, eye-opening methodological aspect: The authors assessed structural brain changes with a longitudinal voxel-based morphometry protocol used in several MR-studies of grey matter plasticity (Boyke, et al., 2008; Draganski, et al., 2004; Draganski, et al., 2006; Driemeyer, et al., 2008; Ilg, et al., 2008).

When the authors performed analyses without proper (i.e., biased) spatial alignment across time points, they found results interpretable as positive effects of intervention. But after unbiased halfway spatial alignment between time-points, the significant effects of training disappeared, yielding a negative result (see Fig. 4, pg. 122). In paper II we tried to minimize bias by implementing longitudinal halfway registration into the TBSS-protocol. By projecting the statistical results back to individual neuroanatomy at each time-point we were able to show, visually, that our protocol yielded excellent anatomical overlap across scans. In paper I we employed a validated longitudinal protocol based on similar principles (Han,

et al., 2006). I believe the present methodological precautions signify a strength of the current thesis.

Another source of variation in the present project, and other work is how we choose to approach the "problem" of multiple comparisons. To a great extent, results in neuroimaging research depend on the method used to adjust for multiple statistical tests. Whole-brain analyses of cognitive intervention is particular harsh in this regard; intervention-costs and logistics limit sample sizes, while the expected effects sizes are often relatively small. This contrasts the hundred thousand voxels of potential analysis in the brain. As a result, the researcher finds her- or him-self on an edge between type I errors on the one side, and type-II errors on the other.

I believe that, although interpretations should be made from multiple comparisons corrected results, later work could benefit from moving towards presenting both corrected and uncorrected data of effect sizes. This holds the potential of giving the reader a broader impression of the data. For instance, in a study of attentional function and cortical thickness the authors found corrected unilateral left-hemisphere associations of an alert attention task and thickness (Westlye, Grydeland, Walhovd, & Fjell, 2010). However, the uncorrected effect-size maps suggested bilateral symmetrical distributed effects (see Figure 1, pg. 6 of the referred paper). These sub-threshold effects failed to reach conventional statistical significance. The effects did, however, point to another story of the relationship between alter attention and the brain. The paper points to how sensitive data interpretation and further research assumptions are on the statistical methods used.

The reviewed issues so far in this section point to important differences in methodological approaches to the study of the human brain plasticity. As far as I can tell, only two independent research groups have used the same MR-protocol for the same intervention. The authors found comparable results in terms of increase grey matter density following training. Nevertheless, the authors failed to find exact anatomical overlap of results (Draganski, et al., 2004; Scholz, et al., 2009). Stronger scientific validity in the field of brain plasticity and longitudinal imaging need to come from standardized measurement and intervention protocols, and further replication of studies—preferably using larger sample sizes. Adherence to common protocol standards is likely to increase the proportion of true research findings in a given field (Ioannidis, 2005).

Optimizing measurement of longitudinal brain structure changes is an ongoing focus for developers of the current MRI analysis suites (e.g., Reuter & Fischl, 2011). Uncertainty about methodological confounders should not block the exiting way of tracking intervention response over time in the living human brain. With these considerations in mind, I believe the present studies serve as pointers to the usefulness of longitudinal imaging in intervention research.

11.5. Final remarks

Today, both the present thesis and other work suggest that humans maintain brain plasticity in response to specific interventions—into old age. We have further shown that cognitive changes following intervention map onto individual brain characteristics; both training-related brain changes (paper I-II), and pre-intervention features of the brain (paper III). The present correlates between brain and cognitive plasticity represents a modest step toward understanding the interrelatedness of these phenomena in aging.

I believe we are seeing a research shift towards studying the extent and limits of brain plasticity in aging. Advanced age confines training effects on cognition (Singer, et al., 2003), and similar constraints have been suggested for the brain (Wagner, et al., 2000). More knowledge is needed on the neurobiological underpinnings of these age-limits, and how they can be modified through proper timing and content of intervention. In the years to come, researchers will likely seek to illuminate the neurobiological and genetic mechanisms through which brain plasticity operates, and how novel neuro-ceutical and gene-therapeutical drugs can modify them. It is unknown how cognitive intervention can be used in combination which pharmaceuticals in order to boost brain plasticity, and at what political and ethical costs (Greely et al., 2008). Still, genetic research is today broadening our understanding of brain plasticity by revealing interactions between training effects and genetic variation (Bellander et al., 2011; Brehmer et al., 2009; Lövdén, Schaefer, Noack, Kanowski, et al., 2011). Such research could further enhance the efficacy of training by taking genetic constraints and facilitators of plasticity into account.

Finally, there is a need to study how various training regimens can be combined to optimize effects on the age trajectories. Randomized controlled trials have documented effects of physical exercise on both brain and behavior (Erickson, et al., 2011); studying the

impact of cognitive and physical exercise combined is one of many exciting research avenues to pursue in the years to come.

12. Conclusions

Conclusions of present findings in order of research questions (pg. 30):

1. Memory training is associated with increased verbal source memory performance, beyond test-retest effects (Paper I-II).

2. Memory training is accompanied by cortical thickening in four brain regions compared with controls. Training effects on thickness are driven by an interaction between cortical thickening in the training group, and thinning in the control group (Paper I).

3. Change in cortical thickness correlates positively with source memory change across groups in the right fusiform and lateral orbitofrontal cortices (Paper I).

4. Longitudinal MD increases are found across groups, with the strongest effects in anterior white matter. Greater MD change correlates with advancing age (Paper II).

5. Training impacts longitudinal FA changes in frontal white matter where MD-change is identified. The effect of training is driven by a significant FA reduction in the control group. In the training group, greater source memory improvement correlates with more positive FA change (Paper II).

6. High participation, low dropout rate, and increased memory performance—beyond test-retest effects—suggest that the memory-training program is feasible for outpatients with subjective memory impairment (paper III, and pg. 58 of this thesis).

7. Greater left, but not right, hippocampal volumes before training is related to better free recall performance following training. The relationship is specific to the left CA2/3 and CA4/DG subfields (Paper III).

8. Depressive symptoms co-vary with recall improvement, indicating larger recall improvement for those with more depressive symptoms at baseline (Paper III).

13. References

Alexander, A. L., Lee, J. E., Lazar, M., & Field, A. S. (2007). Diffusion tensor imaging of the brain. *Neurotherapeutics, 4*(3), 316-329.

Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., . . . for the ACTIVE Study Group. (2002). Effects of Cognitive Training Interventions With Older Adults: A Randomized Controlled Trial. *JAMA, 288*(18), 2271-2281.

Ball, K., Edwards, J. D., & Ross, L. A. (2007). The Impact of Speed of Processing Training on Cognitive and Everyday Functions. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 62*(Special Issue 1), 19-31.

Barrick, T. R., Charlton, R. A., Clark, C. A., & Markus, H. S. (2010). White matter structural decline in normal ageing: A prospective longitudinal study using tract-based spatial statistics. *NeuroImage, 51*(2), 565-577.

Bartzokis, G. (2004). Age-related myelin breakdown: a developmental model of cognitive decline and Alzheimer's disease. *Neurobiology of aging, 25*(1), 5-18; author reply 49-62.

Beaulieu, C. (2009). The biological basis of diffusion anisotropy. In H. Johansen-Berg & T. E. J. Behrens (Eds.), *Diffusion MRI: From Quantitative Measurement to In-vivo Neuroanatomy* (pp. 105-126). London: Elsevier.

Bellander, M., Brehmer, Y., Westerberg, H., Karlsson, S., Fürth, D., Bergman, O., . . . Bäckman, L. (2011). Preliminary evidence that allelic variation in the LMX1A gene influences training-related working memory improvement. *Neuropsychologia, 49*(7), 1938-1942.

Belleville, S., Gilbert, B., Fontaine, F., Gagnon, L., Menard, E., & Gauthier, S. (2006). Improvement of episodic memory in persons with mild cognitive impairment and healthy older adults: evidence from a cognitive intervention program. *Dementia and geriatric cognitive disorders, 22*(5-6), 486-499.

Bengtsson, S. L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., & Ullen, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nat Neurosci, 8*(9), 1148-1150.

Benner, T., Wisco, J. J., van der Kouwe, A. J. W., Fischl, B., Vangel, M. G., Hochberg, F. H., & Sorensen, A. G. (2006). Comparison of Manual and Automatic Section Positioning of Brain MR Images. *Radiology, 239*(1), 246-254.

Bennett, I. J., Madden, D. J., Vaidya, C. J., Howard, D. V., & Howard, J. H., Jr. (2010). Age-related differences in multiple measures of white matter integrity: A diffusion tensor imaging study of healthy aging. *Human brain mapping, 31*(3), 378-390.

Bliss, T. V. P., & Lømo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *The Journal of Physiology, 232*(2), 331-356.

Bower, G. H. (1970). Analysis of a mnemonic device. *American Scientist, 58*(5), 496-510.

Boyke, J., Driemeyer, J., Gaser, C., Buchel, C., & May, A. (2008). Training-Induced Brain Structure Changes in the Elderly. *J. Neurosci., 28*(28), 7031-7035.

Brant-Zawadzki, M., Gillan, G. D., & Nitz, W. R. (1992). MP RAGE: a three-dimensional, T1-weighted, gradient-echo sequence--initial experience in the brain. *Radiology, 182*(3), 769-775.

Brehmer, Y., Westerberg, H., Bellander, M., Fürth, D., Karlsson, S., & Bäckman, L. (2009). Working memory plasticity modulated by dopamine transporter genotype. *Neuroscience Letters, 467*(2), 117-120.

Brenes, G. A. (2003). Cognitive training may improve targeted cognitive functions in older adults. *Evidence-Based Mental Health, 6*(2), 54-54.

Brunborg, H., & Texmon, I. (2010). Befolkningsframskrivninger 2010-2060. *Økonomiske analyser*(4), 28-39.

Brækhus, A., Ulstein, I., Wyller, T. B., & Engedal, K. (2011). The Memory Clinic - outpatient assessment when dementia is suspected. *Tidsskrift for den Norske lægeforening, 131*(22), 2254-2257.

Buonomano, D. V., & Merzenich, M. M. (1998). CORTICAL PLASTICITY: From Synapses to Maps. *Annual Review of Neuroscience, 21*(1), 149-186.

Burke, S. N., & Barnes, C. A. (2006). Neural plasticity in the ageing brain. *Nature Reviews Neuroscience, 7*(1), 30-40.

Burns, A., & Iliffe, S. (2009). Dementia. *BMJ, 338*, b75.

Buschert, V., Bokde, A. L., & Hampel, H. (2010). Cognitive intervention in Alzheimer disease. *Nature reviews. Neurology, 6*(9), 508-517.

Byers, A. L., & Yaffe, K. (2011). Depression and risk of developing dementia. *Nature reviews. Neurology*.

Cabeza, R., Nyberg, L., & Park, D. C. (2005). Cognitive neuroscience of aging: Emergence of a new discipline. In R. Cabeza, L. Nyberg & D. C. Park (Eds.), *Cognitive neuroscience of aging : linking cognitive and cerebral aging* (pp. 3-15). Oxford ; New York: Oxford University Press.

Ceccarelli, A., Rocca, M. A., Pagani, E., Falini, A., Comi, G., & Filippi, M. (2009). Cognitive learning is associated with gray matter changes in healthy human individuals: A tensor-based morphometry study. *NeuroImage, 48*(3), 585-589.

Chanraud, S., Zahr, N., Sullivan, E., & Pfefferbaum, A. (2010). MR Diffusion Tensor Imaging: A Window into White Matter Integrity of the Working Brain. *Neuropsychology Review, 20*(2), 209-225.

Chklovskii, D. B. (2004). Synaptic Connectivity and Neuronal Morphology: Two Sides of the Same Coin. *Neuron, 43*(5), 609-617.

Chklovskii, D. B., Mel, B. W., & Svoboda, K. (2004). Cortical rewiring and information storage. *Nature, 431*(7010), 782-788.

Cicero, M. T. *De Oratore* (E. W. Sutton & H. Rackham, Trans.). Cambridge, MA: Harvard University Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J. ; Hove: Erlbaum Associates.

Concha, L., Livy, D. J., Beaulieu, C., Wheatley, B. M., & Gross, D. W. (2010). In Vivo Diffusion Tensor Imaging and Histopathology of the Fimbria-Fornix in Temporal Lobe Epilepsy. *J. Neurosci., 30*(3), 996-1002.

Cramer, S. C., Sur, M., Dobkin, B. H., O'Brien, C., Sanger, T. D., Trojanowski, J. Q., . . . Vinogradov, S. (2011). Harnessing neuroplasticity for clinical applications. *Brain, 134*(Pt 6), 1591-1609.

Cutler, N. E., & Whitelaw, N. A. (2005). American Perceptions of Aging in the 21st Century [APA21], 2000: Inter-university Consortium for Political and Social Research (ICPSR).

Davis, S. W., Dennis, N. A., Buchler, N. G., White, L. E., Madden, D. J., & Cabeza, R. (2009). Assessing the effects of age on long white matter tracts using diffusion tensor tractography. *NeuroImage, 46*(2), 530-541.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000a). California Verbal Learning Test-Second Edition (CVLT-II). *San Antonio: The Psychological Corporation*.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000b). California Verbal Learning Test-Second Edition (CVLT-II) Manual. *San Antonio: The Psychological Corporation*.

Demerens, C., Stankoff, B., Logak, M., Anglade, P., Allinquant, B., Couraud, F., . . . Lubetzki, C. (1996). Induction of myelination in the central nervous system by electrical activity. *Proceedings of the National Academy of Sciences of the United States of America, 93*(18), 9887-9892.

Dickstein, D. L., Morrison, J. H., & Hof, P. R. (2009). Neuropathology of Aging. In W. Jagust & M. D'Esposito (Eds.), *Imaging the Aging Brain* (1st ed., pp. 27-40): Oxford University Press.

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004). Neuroplasticity: Changes in grey matter induced by training. *Nature, 427*(6972), 311-312.

Draganski, B., Gaser, C., Kempermann, G., Kuhn, H. G., Winkler, J., Buchel, C., & May, A. (2006). Temporal and Spatial Dynamics of Brain Structure Changes during Extensive Learning. *The Journal of neuroscience, 26*(23), 6314-6317.

Draganski, B., & May, A. (2008). Training-induced structural changes in the adult human brain. *Behavioural Brain Research, 192*(1), 137-142.

Driemeyer, J., Boyke, J., Gaser, C., Büchel, C., & May, A. (2008). Changes in gray matter induced by learning--revisited. *PLoS ONE, 3*(7).

Duan, H., Wearne, S. L., Rocher, A. B., Macedo, A., Morrison, J. H., & Hof, P. R. (2003). Age-related dendritic and spine changes in corticocortically projecting neurons in macaque monkeys. *Cerebral cortex, 13*(9), 950-961.

Erickson, K. I., Boot, W. R., Basak, C., Neider, M. B., Prakash, R. S., Voss, M. W., . . . Kramer, A. F. (2010). Striatal volume predicts level of video game skill acquisition. *Cerebral cortex, 20*(11), 2522-2530.

Erickson, K. I., Voss, M. W., Prakash, R. S., Basak, C., Szabo, A., Chaddock, L., . . . Kramer, A. F. (2011). Exercise training increases size of hippocampus and improves memory. *Proceedings of the National Academy of Sciences of the United States of America, 108*(7), 3017-3022.

Erk, S., Spottke, A., Meisen, A., Wagner, M., Walter, H., & Jessen, F. (2011). Evidence of neuronal compensation during episodic memory in subjective memory impairment. *Archives of general psychiatry, 68*(8), 845-852.

Esiri, M. M. (2007). Ageing and the brain. *The Journal of pathology, 211*(2), 181-187.

Espeseth, T., Westlye, L. T., Fjell, A. M., Walhovd, K. B., Rootwelt, H., & Reinvang, I. (2008). Accelerated age-related cortical thinning in healthy carriers of apolipoprotein E epsilon 4. *Neurobiology of aging, 29*(3), 329-340.

Fernandez, A. (2010). Transforming Brain Health with Digital Tools to Assess, Enhance and Treat Cognition across the Lifespan *The State of the Brain Fitness Market 2010*

Fields, R. D. (2008). White matter in learning, cognition and psychiatric disorders. *Trends in Neurosciences, 31*(7), 361-370.

Finch, C. E. (2009). The neurobiology of middle-age has arrived. *Neurobiology of aging, 30*(4), 515-520.

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A, 97*(20), 11050-11055.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., . . . Dale, A. M. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron, 33*(3), 341-355.

Fjell, A. M., & Walhovd, K. B. (2010). Structural brain changes in aging: courses, causes and cognitive consequences. *Reviews in the neurosciences, 21*(3), 187-221.

Fjell, A. M., Walhovd, K. B., Fennema-Notestine, C., McEvoy, L. K., Hagler, D. J., Holland, D., . . . Dale, A. M. (2009). One-year brain atrophy evident in healthy aging. *The Journal of neuroscience, 29*(48), 15223-15231.

Fjell, A. M., Westlye, L. T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., . . . Walhovd, K. B. (2009). High consistency of regional cortical thinning in aging across multiple samples. *Cerebral cortex, 19*(9), 2001-2012.

Fotenos, A. F., Snyder, A. Z., Girton, L. E., Morris, J. C., & Buckner, R. L. (2005). Normative estimates of cross-sectional and longitudinal brain volume decline in aging and AD. *Neurology, 64*(6), 1032-1039.

Giorgio, A., Santelli, L., Tomassini, V., Bosnell, R., Smith, S. M., De Stefano, N., & Johansen-Berg, H. (2010). Age-related changes in grey and white matter structure throughout adulthood. *NeuroImage, 51*(3), 943-951.

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N., Friston, K. J., & Frackowiak, R. S. (2001). A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage, 14*(1 Pt 1), 21-36.

Greely, H., Sahakian, B., Harris, J., Kessler, R. C., Gazzaniga, M., Campbell, P., & Farah, M. J. (2008). Towards responsible use of cognitive-enhancing drugs by the healthy. *Nature, 456*(7223), 702-705.

Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., . . . Fischl, B. (2006). Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *NeuroImage, 32*(1), 180-194.

Hanggi, J., Koeneke, S., Bezzola, L., & Jäncke, L. (2010). Structural neuroplasticity in the sensorimotor network of professional female ballet dancers. *Hum Brain Mapp, 31*(8), 1196-1206.

Hashemi, R. H., Bradley, W. G., & Lisanti, C. J. (2010). *MRI: The Basics*: Lippincott Williams & Wilkins.

Head, D., Buckner, R. L., Shimony, J. S., Williams, L. E., Akbudak, E., Conturo, T. E., . . . Snyder, A. Z. (2004). Differential Vulnerability of Anterior White Matter in Nondemented Aging with Minimal Acceleration in Dementia of the Alzheimer Type: Evidence from Diffusion Tensor Imaging. *Cerebral Cortex, 14*(4), 410-423.

Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the ageing mind: a view from cognitive neuroscience. *Nature Reviews Neuroscience, 5*(2), 87-96.

Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberger, U. (2008). Enrichment Effects on Adult Cognitive Development. *Psychological Science in the Public Interest, 9*(1), 1-65.

Holtmaat, A., & Svoboda, K. (2009). Experience-dependent structural synaptic plasticity in the mammalian brain. *Nat Rev Neurosci, 10*(9), 647-658.

Holtmaat, A., Wilbrecht, L., Knott, G. W., Welker, E., & Svoboda, K. (2006). Experience-dependent and cell-type-specific spine growth in the neocortex. *Nature, 441*(7096), 979-983.

Ilg, R., Wohlschlager, A. M., Gaser, C., Liebau, Y., Dauner, R., Woller, A., . . . Muhlau, M. (2008). Gray Matter Increase Induced by Practice Correlates with Task-Specific Activation: A Combined Functional and Morphometric Magnetic Resonance Imaging Study. *The Journal of neuroscience, 28*(16), 4210-4215.

Iliffe, S., & Pealing, L. (2010). Subjective memory problems. *BMJ, 340*, c1425.

Imfeld, A., Oechslin, M. S., Meyer, M., Loenneker, T., & Jäncke, L. (2009). White matter plasticity in the corticospinal tract of musicians: a diffusion tensor imaging study. *NeuroImage, 46*(3), 600-607.

Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLoS Medicine, 2*(8), e124.

Ishibashi, T., Dakin, K. A., Stevens, B., Lee, P. R., Kozlov, S. V., Stewart, C. L., & Fields, R. D. (2006). Astrocytes Promote Myelination in Response to Electrical Impulses. *Neuron, 49*(6), 823-832.

Jbabdi, S., Behrens, T. E. J., & Smith, S. M. (2010). Crossing fibres in tract-based spatial statistics. *NeuroImage, 49*(1), 249-256.

Jenkins, W. M., Merzenich, M. M., Ochs, M. T., Allard, T., & Guic-Robles, E. (1990). Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. *J Neurophysiol, 63*(1), 82-104.

Jessen, F., Wiese, B., Bachmann, C., Eifflaender-Gorfer, S., Haller, F., Kolsch, H., . . . Bickel, H. (2010). Prediction of dementia by subjective memory impairment: effects of severity and temporal association with cognitive impairment. *Archives of general psychiatry, 67*(4), 414-422.

Johansen-Berg, H. (2011). The future of functionally-related structural change assessment. *NeuroImage*.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3-28.

Jones, S., Nyberg, L., Sandblom, J., Stigsdotter Neely, A., Ingvar, M., Magnus Petersson, K., & Bäckman, L. (2006). Cognitive and neural plasticity in aging: General and task-specific limitations. *Neuroscience & Biobehavioral Reviews, 30*(6), 864-871.

Jonker, C., Geerlings, M. I., & Schmand, B. (2000). Are memory complaints predictive for dementia? A review of clinical and population-based studies. *International Journal of Geriatric Psychiatry, 15*(11), 983-991.

Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., . . . Dale, A. (2006). Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. *NeuroImage, 30*(2), 436-443.

Jovicich, J., Czanner, S., Han, X., Salat, D., van der Kouwe, A., Quinn, B., . . . Fischl, B. (2009). MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: Reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. *NeuroImage, 46*(1), 177-192.

Jäncke, L. (2009). The plastic human brain. *Restorative neurology and neuroscience, 27*(5), 521-538.

Kaes, T. (1907). *Die Grosshirnrinde des Menschen in ihren Massen und in ihrem Fasergehalt: Ein gehirnanatomischer Atlas mit erläuterndem Text und schematischer Zeichnung*. Jena, Germany: Fischer.

Keller, T. A., & Just, M. A. (2009). Altering cortical connectivity: remediation-induced changes in the white matter of poor readers. *Neuron, 64*(5), 624-631.

Kennedy, K. M., & Raz, N. (2009). Aging white matter and cognition: differential effects of regional variations in diffusion properties on memory, executive functions, and speed. *Neuropsychologia, 47*(3), 916-927.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlstrom, K., . . . Westerberg, H. (2005). Computerized training of working memory in children with ADHD--a randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*(2), 177-186.

Kramer, J. H., Blusewicz, M. J., & Preston, K. A. (1989). The premature aging hypothesis: old before its time? *Journal of consulting and clinical psychology, 57*(2), 257-262.

Kuperberg, G. R., Broome, M. R., McGuire, P. K., David, A. S., Eddy, M., Ozawa, F., . . . Fischl, B. (2003). Regionally localized thinning of the cerebral cortex in schizophrenia. *Archives of general psychiatry, 60*(9), 878-888.

Kwok, V., Niu, Z., Kay, P., Zhou, K., Mo, L., Jin, Z., . . . Tan, L. H. (2011). Learning new color names produces rapid increase in gray matter in the intact adult human cortex. *Proceedings of the National Academy of Sciences of the United States of America, 108*(16), 6686-6688.

Le Bihan, D. (2003). Looking into the functional architecture of the brain with diffusion MRI. *Nat Neurosci, 4*(6), 469-480.

Lerch, J. P., Yiu, A. P., Martinez-Canabal, A., Pekar, T., Bohbot, V. D., Frankland, P. W., . . . Sled, J. G. (2011). Maze training in mice induces MRI-detectable brain shape changes specific to the type of learning. *NeuroImage, 54*(3), 2086-2095.

Lindenberger, U., von Oertzen, T., Ghisletta, P., & Hertzog, C. (2011). Cross-sectional age variance extraction: what's change got to do with it? *Psychology and aging, 26*(1), 34-47.

Luscher, C., Nicoll, R. A., Malenka, R. C., & Muller, D. (2000). Synaptic plasticity and dynamic modulation of the postsynaptic membrane. *Nat Neurosci, 3*(6), 545-550.

Lømo, T. (1966). Frequency potentiation of excitatory synaptic activity in the dentate area of the hippocampal formation. *Acta physiol. scand., 68*(suppl 277), 128.

Lövdén, M., Bodammer, N. C., Kühn, S., Kaufmann, J., Schütze, H., Tempelmann, C., . . . Lindenberger, U. (2010). Experience-dependent plasticity of white-matter microstructure extends into old age. *Neuropsychologia, 48*(13), 3878-3883.

Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin, 136*(4), 659-676.

Lövdén, M., Schaefer, S., Noack, H., Bodammer, N. C., Kuhn, S., Heinze, H. J., . . . Lindenberger, U. (2011). Spatial navigation training protects the hippocampus against age-related changes during early and late adulthood. *Neurobiology of aging*.

Lövdén, M., Schaefer, S., Noack, H., Kanowski, M., Kaufmann, J., Tempelmann, C., . . . Bäckman, L. (2011). Performance-related increases in hippocampal N-acetylaspartate (NAA) induced by spatial navigation training are restricted to BDNF Val homozygotes. *Cerebral cortex, 21*(6), 1435-1442.

Madden, D. J., Bennett, I. J., Burzynska, A., Potter, G. G., Chen, N.-k., & Song, A. W. (2011). Diffusion tensor imaging of cerebral white matter integrity in cognitive aging. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., & Frith, C. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences of the United States of America, 8*(97), 4398–4403.

Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: A structural MRI and neuropsychological analysis. *Hippocampus, 16*(12), 1091-1101.

Mahncke, H. W., Bronstone, A., Merzenich, M. M., & Aage, R. M. (2006). Brain plasticity and functional losses in the aged: scientific bases for a novel intervention *Progress in Brain Research* (Vol. Volume 157, pp. 81-109): Elsevier.

Marner, L., Nyengaard, J. R., Tang, Y., & Pakkenberg, B. (2003). Marked loss of myelinated nerve fibers in the human brain with age. *The Journal of comparative neurology, 462*(2), 144-152.

May, A., & Gaser, C. (2006). Magnetic resonance-based morphometry: a window into structural plasticity of the brain. *Curr Opin Neurol, 19*(4), 407-411.

Murphy, E. A., Holland, D., Donohue, M., McEvoy, L. K., Hagler Jr, D. J., Dale, A. M., & Brewer, J. B. (2010). Six-month atrophy in MTL structures is associated with subsequent memory decline in elderly controls. *NeuroImage, 53*(4), 1310-1317.

Nair, G., Tanahashi, Y., Low, H. P., Billings-Gagliardi, S., Schwartz, W. J., & Duong, T. Q. (2005). Myelination and long diffusion times alter diffusion-tensor-imaging contrast in myelin-deficient shiverer mice. *NeuroImage, 28*(1), 165-174.

Naveh-Benjamin, M., & Old, S. R. (2008). 2.41 - Aging and Memory. In H. B. Editor-in-Chief:¬†¬†John (Ed.), *Learning and Memory: A Comprehensive Reference* (pp. 787-808). Oxford: Academic Press.

Neely, A. S., & Backman, L. (1993). MAINTENANCE OF GAINS FOLLOWING MULTIFACTORIAL AND UNIFACTORIAL MEMORY TRAINING IN LATE ADULTHOOD. *Educational Gerontology, 19*(2), 105 - 117.

Nilsson, L. G., Sternang, O., Ronnlund, M., & Nyberg, L. (2009). Challenging the notion of an early-onset of cognitive decline. *Neurobiology of aging, 30*(4), 521-524; discussion 530-523.

O'Sullivan, M., Jones, D. K., Summers, P. E., Morris, R. G., Williams, S. C., & Markus, H. S. (2001). Evidence for cortical "disconnection" as a mechanism of age-related cognitive decline. *Neurology, 57*(4), 632-638.

Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S., Burns, A. S., . . . Ballard, C. G. (2010). Putting brain training to the test. *Nature, 465*(7299), 775-778.

Pakkenberg, B., & Gundersen, H. J. (1997). Neocortical neuron number in humans: effect of sex and age. *The Journal of comparative neurology, 384*(2), 312-320.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and aging, 17*(2), 299-320.

Park, D. C., & Reuter-Lorenz, P. (2009). The Adaptive Brain: Aging and Neurocognitive Scaffolding. *Annual Review of Psychology, 60*(1), 173-196.

Pascual-Leone, A., Amedi, A., Fregni, F., & Merabet, L. B. (2005). The Plastic Human Brain Cortex. *Annual Review of Neuroscience, 28*(1), 377-401.

Peters, A. (2002). The effects of normal aging on myelin and nerve fibers: a review. *Journal of neurocytology, 31*(8-9), 581-593.

Petrella, J. R., Mattay, V. S., & Doraiswamy, P. M. (2008). Imaging genetics of brain longevity and mental wellness: the next frontier? *Radiology, 246*(1), 20-32.

Pierpaoli, C., Barnett, A., Pajevic, S., Chen, R., Penix, L. R., Virta, A., & Basser, P. (2001). Water diffusion changes in Wallerian degeneration and their dependence on white matter architecture. *Neuroimage, 13*(6 Pt 1), 1174-1185.

Quallo, M. M., Price, C. J., Ueno, K., Asamizuya, T., Cheng, K., Lemon, R. N., & Iriki, A. (2009). Gray and white matter changes associated with tool-use learning in macaque monkeys. *Proceedings of the National Academy of Sciences, 106*(43), 18379-18384.

Ranganath, C., Flegal, Kristin E., & Kelly, Laura L. (2011). Can Cognitive Training Improve Episodic Memory? *Neuron, 72*(5), 688-691.

Raz, N., & Lindenberger, U. (2011). Only time will tell: cross-sectional studies offer no solution to the age-brain-cognition triangle: comment on Salthouse (2011). *Psychological Bulletin, 137*(5), 790-795.

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., . . . Acker, J. D. (2005). Regional Brain Changes in Aging Healthy Adults: General Trends, Individual Differences and Modifiers. *Cerebral cortex, 15*(11), 1676-1689.

Raz, N., Rodrigue, K. M., Head, D., Kennedy, K. M., & Acker, J. D. (2004). Differential aging of the medial temporal lobe: a study of a five-year change. *Neurology, 62*(3), 433-438.

Rebok, G. W., Carlson, M. C., & Langbaum, J. B. S. (2007). Training and Maintaining Memory Abilities in Healthy Older Adults: Traditional and Novel Approaches. *J Gerontol B Psychol Sci Soc Sci, 62*(suppl_Special_Issue_1), 53-61.

Reichman, W. E., Fiocco, A. J., & Rose, N. S. (2010). Exercising the brain to avoid cognitive decline: examining the evidence. *Aging Health, 6*(5), 565-584.

Reid, L., & MacLullich, A. (2006). Subjective Memory Complaints and Cognitive Impairment in Older People. *Dement Geriatr Cogn Disord, 22*(5-6), 471-485.

Reisberg, B., Shulman, M. B., Torossian, C., Leng, L., & Zhu, W. (2010). Outcome over seven years of healthy adults with and without subjective cognitive impairment. *Alzheimer's and Dementia, 6*(1), 11-24.

Reuter, M., & Fischl, B. (2011). Avoiding asymmetry-induced bias in longitudinal image processing. *NeuroImage, 57*(1), 19-21.

Rosas, H. D., Liu, A. K., Hersch, S., Glessner, M., Ferrante, R. J., Salat, D. H., . . . Fischl, B. (2002). Regional and progressive thinning of the cortical ribbon in Huntington's disease. *Neurology, 58*(5), 695-701.

Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S., Busa, E., . . . Fischl, B. (2004). Thinning of the cerebral cortex in aging. *Cerebral cortex, 14*(7), 721-730.

Salat, D. H., Tuch, D. S., Greve, D. N., van der Kouwe, A. J. W., Hevelone, N. D., Zaleta, A. K., . . . Dale, A. M. (2005). Age-related alterations in white matter microstructure measured by diffusion tensor imaging. *Neurobiology of Aging, 26*(8), 1215-1227.

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of aging, 30*(4), 507-514.

Salthouse, T. A. (2011). Neuroanatomical substrates of age-related cognitive decline. *Psychological Bulletin, 137*(5), 753-784.

Saykin, A. J., Wishart, H. A., Rabin, L. A., Santulli, R. B., Flashman, L. A., West, J. D., . . . Mamourian, A. C. (2006). Older adults with cognitive complaints show brain atrophy similar to that of amnestic MCI. *Neurology, 67*(5), 834-842.

Scargle, J. D. (2000). Publication Bias: The “File-Drawer” Problem in Scientific Inference. *Journal of Scientific Exploration, 14*(1), 91-106.

Schaie, K. W. (2009). "When does age-related cognitive decline begin?" Salthouse again reifies the "cross-sectional fallacy". *Neurobiology of aging, 30*(4), 528-529.

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia, 33*(8), 1047-1055.

Schlaug, G., Jäncke, L., Huang, Y., & Steinmetz, H. (1995). In vivo evidence of structural brain asymmetry in musicians. *Science, 267*(5198), 699-701.

Schmierer, K., Wheeler-Kingshott, C. A., Tozer, D. J., Boulby, P. A., Parkes, H. G., Yousry, T. A., . . . Miller, D. H. (2008). Quantitative magnetic resonance of postmortem multiple sclerosis brain before and after fixation. *Magnetic resonance in medicine, 59*(2), 268-277.

Scholz, J., Klein, M. C., Behrens, T. E. J., & Johansen-Berg, H. (2009). Training induces changes in white-matter architecture. *Nat Neurosci, 12*(11), 1370-1371.

Singer, T., Lindenberger, U., & Baltes, P. B. (2003). Plasticity of memory for new learning in very old age: a story of major loss? *Psychology and aging, 18*(2), 306-317.

Singh-Manoux, A., Kivimaki, M., Glymour, M. M., Elbaz, A., Berr, C., Ebmeier, K. P., . . . Dugravot, A. (2012). Timing of onset of cognitive decline: results from Whitehall II prospective cohort study. *BMJ, 344*.

Smith, D. E., Rapp, P. R., McKay, H. M., Roberts, J. A., & Tuszynski, M. H. (2004). Memory impairment in aged primates is associated with focal death of cortical

neurons and atrophy of subcortical neurons. *The Journal of neuroscience, 24*(18), 4373-4381.

Smith, G. E., Housen, P., Yaffe, K., Ruff, R., Kennison, R. F., Mahncke, H. W., & Zelinski, E. M. (2009). A Cognitive Training Program Based on Principles of Brain Plasticity: Results from the Improvement in Memory with Plasticity-based Adaptive Cognitive Training (IMPACT) Study. *Journal of the American Geriatrics Society*.

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., . . . Behrens, T. E. J. (2006). Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage, 31*(4), 1487-1505.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., . . . Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage, 23 Suppl 1*, S208-219.

Smith, S. M., Johansen-Berg, H., Jenkinson, M., Rueckert, D., Nichols, T. E., Miller, K. L., . . . Behrens, T. E. (2007). Acquisition and voxelwise analysis of multi-subject diffusion data with tract-based spatial statistics. *Nature protocols, 2*(3), 499-503.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage, 44*(1), 83-98.

Song, S. K., Sun, S. W., Ramsbottom, M. J., Chang, C., Russell, J., & Cross, A. H. (2002). Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *Neuroimage, 17*(3), 1429-1436.

Squire, L. R., Stark, C. E., & Clark, R. E. (2004). The medial temporal lobe. *Annual review of neuroscience, 27*, 279-306.

Striepens, N., Scheef, L., Wind, A., Popp, J., Spottke, A., Cooper-Mahkorn, D., . . . Jessen, F. (2010). Volume loss of the medial temporal lobe structures in subjective memory impairment. *Dementia and geriatric cognitive disorders, 29*(1), 75-81.

Sullivan, E. V., Marsh, L., & Pfefferbaum, A. (2005). Preservation of hippocampal volume throughout adulthood in healthy men and women. *Neurobiology of aging, 26*(7), 1093-1098.

Sullivan, E. V., & Pfefferbaum, A. (2006). Diffusion tensor imaging and aging. *Neurosci Biobehav Rev, 30*(6), 749-761.

Sullivan, E. V., Pfefferbaum, A., Swan, G. E., & Carmelli, D. (2001). Heritability of hippocampal size in elderly twin men: equivalent influence from genes and environment. *Hippocampus, 11*(6), 754-762.

Takeuchi, H., Sekiguchi, A., Taki, Y., Yokoyama, S., Yomogida, Y., Komuro, N., . . . Kawashima, R. (2010). Training of Working Memory Impacts Structural Connectivity. *J. Neurosci., 30*(9), 3297-3303.

Tang, Y. Y., Lu, Q., Geng, X., Stein, E. A., Yang, Y., & Posner, M. I. (2010). Short-term meditation induces white matter changes in the anterior cingulate. *Proceedings of the National Academy of Sciences of the United States of America, 107*(35), 15649-15652.

Taubert, M., Draganski, B., Anwander, A., Muller, K., Horstmann, A., Villringer, A., & Ragert, P. (2010). Dynamic Properties of Human Brain Structure: Learning-Related Changes in Cortical Areas and Associated Fiber Connections. *J. Neurosci., 30*(35), 11670-11677.

Thomas, A. G., Marrett, S., Saad, Z. S., Ruff, D. A., Martin, A., & Bandettini, P. A. (2009). Functional but not structural changes associated with learning: an exploration of longitudinal voxel-based morphometry (VBM). *NeuroImage, 48*(1), 117-125.

Toga, A. W., Thompson, P. M., & Sowell, E. R. (2006). Mapping brain maturation. *Trends in Neurosciences, 29*(3), 148-159.

Tuch, D. S., Reese, T. G., Wiegell, M. R., & Wedeen, V. J. (2003). Diffusion MRI of complex neural architecture. *Neuron, 40*(5), 885-895.

Tupler, L. A., Krishnan, K. R., Greenberg, D. L., Marcovina, S. M., Payne, M. E., MacFall, J. R., . . . Doraiswamy, P. M. (2007). Predicting memory decline in normal elderly: genetics, MRI, and cognitive reserve. *Neurobiology of aging, 28*(11), 1644-1656.

Tyszka, J. M., Readhead, C., Bearer, E. L., Pautler, R. G., & Jacobs, R. E. (2006). Statistical diffusion tensor histology reveals regional dysmyelination effects in the shiverer mouse mutant. *NeuroImage, 29*(4), 1058-1065.

van der Kouwe, A. J., Benner, T., Fischl, B., Schmitt, F., Salat, D. H., Harder, M., . . . Dale, A. M. (2005). On-line automatic slice positioning for brain MR imaging. *Neuroimage, 27*(1), 222-230.

Van Leemput, K., Bakkour, A., Benner, T., Wiggins, G., Wald, L. L., Augustinack, J., . . . Fischl, B. (2009). Automated segmentation of hippocampal subfields from ultra-high resolution in vivo MRI. *Hippocampus, 19*(6), 549-557.

Verhaeghen, P., & Marcoen, A. (1994). Production deficiency hypothesis revisited: Adult age differences in strategy use as a function of processing resources. *Aging, Neuropsychology, and Cognition, 1*(4), 323-338.

Verhaeghen, P., & Marcoen, A. (1996). On the mechanisms of plasticity in young and older adults after instruction in the method of loci: evidence for an amplification model. *Psychology and aging, 11*(1), 164-178.

Verhaeghen, P., Marcoen, A., & Goossens, L. (1992). Improving Memory Performance in the Aged Through Mnemonic Training: A Meta-Analytic Study. *Psychology and Aging, 7*(2), 242-251.

Wagner, A. P., Schmoll, H., Badan, I., Platt, D., & Kessler, C. (2000). Brain plasticity: to what extent do aged animals retain the capacity to coordinate gene activity in response to acute challenges. *Experimental Gerontology, 35*(9-10), 1211-1227.

Wake, H., Lee, P. R., & Fields, R. D. (2011). Control of Local Protein Synthesis and Initial Events in Myelination by Action Potentials. *Science, 333*(6049), 1647-1651.

Walhovd, K. B., Westlye, L. T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., . . . Fjell, A. M. (2011). Consistent neuroanatomical age-related volume differences across multiple samples. *Neurobiology of aging, 32*(5), 916-932.

Walhovd, K. B., Westlye, L. T., Moe, V., Slinning, K., Due-Tonnessen, P., Bjornerud, A., . . . Fjell, A. M. (2010). White matter characteristics and cognition in prenatally opiate- and polysubstance-exposed children: a diffusion tensor imaging study. *AJNR. American journal of neuroradiology, 31*(5), 894-900.

Weishaupt, D., Köchli, V. D., & Marinček, B. (2006). *How does MRI work?: an introduction to the physics and function of magnetic resonance imaging*: Springer.

Wen, W., Sachdev, P. S., Li, J. J., Chen, X., & Anstey, K. J. (2009). White matter hyperintensities in the forties: their prevalence and topography in an epidemiological sample aged 44-48. *Human brain mapping, 30*(4), 1155-1167.

Wenger, E., Schaefer, S., Noack, H., Kuhn, S., Martensson, J., Heinze, H. J., . . . Lövdén, M. (2012). Cortical thickness changes following spatial navigation training in adulthood and aging. *NeuroImage, 59*, 3389–3397.

West, M. J., Coleman, P. D., Flood, D. G., & Troncoso, J. C. (1994). Differences in the pattern of hippocampal neuronal loss in normal ageing and Alzheimer's disease. *Lancet, 344*(8925), 769-772.

Westlye, L. T., Grydeland, H., Walhovd, K. B., & Fjell, A. M. (2010). Associations between Regional Cortical Thickness and Attentional Networks as Measured by the Attention Network Test. *Cerebral cortex, 21*(2), 345-356.

Westlye, L. T., Walhovd, K. B., Dale, A. M., Bjornerud, A., Due-Tonnessen, P., Engvig, A., . . . Fjell, A. M. (2010). Life-span changes of the human brain White matter: diffusion tensor imaging (DTI) and volumetry. *Cerebral cortex, 20*(9), 2055-2068.

Wiegell, M. R., Larsson, H. B., & Wedeen, V. J. (2000). Fiber crossing in human brain depicted with diffusion tensor MR imaging. *Radiology, 217*(3), 897-903.

Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., . . . for the ACTIVE Study Group. (2006). Long-term Effects of Cognitive Training on Everyday Functional Outcomes in Older Adults. *JAMA, 296*(23), 2805-2814.

Wolinsky, F. D., Vander Weg, M. W., Martin, R., Unverzagt, F. W., Ball, K., Jones, R. N., & Tennstedt, S. L. (2009). The Effect of Speed-of-Processing Training on Depressive Symptoms in ACTIVE. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 64A*(4), 468-472.

Woollett, K., & Maguire, E. A. (2011). Acquiring "the Knowledge" of London's Layout Drives Structural Brain Changes. *Current biology, 21*(24), 2109-2114.

Xu, T., Yu, X., Perlik, A. J., Tobin, W. F., Zweig, J. A., Tennant, K., . . . Zuo, Y. (2009). Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature, 462*(7275), 915-919.

Yang, G., Pan, F., & Gan, W.-B. (2009). Stably maintained dendritic spines are associated with lifelong memories. *Nature, 462*(7275), 920-924.

Yasumatsu, N., Matsuzaki, M., Miyazaki, T., Noguchi, J., & Kasai, H. (2008). Principles of Long-Term Dynamics of Dendritic Spines. *J. Neurosci., 28*(50), 13592-13608.

Yates, F. A. (1966). *The art of memory*. Chicago,: University of Chicago Press.

Ziegler, G., Dahnke, R., Jäncke, L., Yotter, R. A., May, A., & Gaser, C. (2011). Brain structural trajectories over the adult lifespan. *Human brain mapping*.

Østby, Y., Tamnes, C. K., Fjell, A. M., Westlye, L. T., Due-Tønnessen, P., & Walhovd, K. B. (2009). Heterogeneity in subcortical brain development: A structural magnetic resonance imaging study of brain maturation from 8 to 30 years. *The Journal of neuroscience, 29*(38), 11772-11782.

Appendix

In-depth description of weekly memory training curriculum

Session 1

In session 1, the instructor gave an overview of the eight weeks and presented the goals of the program. Participants were given a personal calendar that contained the training schedule of the intervention. The calendar included space for marking completed class-sessions and homework assignment for participants to track their own progress. Participants were also given a personal 12-page memory handbook describing what memory is (definitions, the modal model), practical memory principles (attention, association, visulization, repetition, and spatial arrangement), and a review of the use of MoL. Next, the instructor held a talk with three topics:

1. A basic review of memory processes focusing on the 3-step modal model – encoding, storage and retrieval. This model was referred to throughout the program (e.g. when addressing the importance of attention for successful remembering).

2. Memory and aging. The aim was to increase awareness that modifiable factors exist that are likely to impact individual differences in aging – including physical activity, sleep, mood, and cognitive exercise.

3. A visual slide demonstration of the MoL.

Instruction in the Method of loci

The MoL is a visual mnemonic technique known to be applicable to older populations. However, as older adults tend to have difficulty generating visual images and visual imagery associations, the first and second week of the program paid particular attention to visual imagery and associations. At the end of the first session, participants were demonstrated the use of MoL using a 3-step visual slide example. See supplemental document 1 of paper III for a visual example of the demonstration.

Step 1 - acquisition of loci-route

The basis for MoL is to visualize a series of mental landmarks or loci (e.g., various rooms in one's house). These loci make up a route—the loci-route. To illustrate how a loci-route can look like, participants were first shown photos of a 10-point loci-route in a person´s house. The instructor emphasized that the loci should be within a geographically limited area where the participant had been several times before (e.g. the house where one grew up). The loci had to be continuous (follow each other in space), and to constitute specific, fixed places, such as "entrance", "hallway", "kitchen", and "balcony". Participants were told that they could walk through walls or floors in their minds eye – if needed to successfully complete their route.

Steps 2-3 – visualize to-be-remembered information at each locus

After the loci-route is established, the second step in using the MoL is to link to-be-remembered bits of information as images (or visual associations if the information is abstract) at each locus. This is the encoding phase. The third step involves revisiting the landmarks mentally to recollect the information associated with each locus. This is the recall phase.

Steps 1-3 were demonstrated by showing pictures of various concrete words (fire hydrant, helium balloon, berries, ox, etc.) combined with each of the 10 loci in the route. One picture was shown at each locus.

Finally, after the demonstration was over, the participants were told that fire **hydr**ant was a visual association for **hydr**ogen, helium balloon an association for helium, and that they had actually been shown a way to learn abstract verbal information (part of the periodic table of the elements) through use of visual associations. This was done to increase participant motivation and to demonstrate the purpose of using visual images in cases where the to-be remembered information is hard to visualize in it self.

Before the end of the session, participants were handed out four homework assignments focusing on creating and learning to use their own loci-routes at their own pace at home. In the fourth assignment, participants were instructed to expand their loci-route (or create a new one) to include 16 loci.

Session 2

This session was dedicated for troubleshooting and learning to use the MoL more efficiently. First, participants could troubleshoot the use of the MoL with the instructor and discuss the homework assignments. Second, the instructor recapitulated on the MoL and explained that the MoL is like a library to store and keep information that one needs to remember organized. Third, the instructor told the participants that to improve the use of MoL, they should start visualization of two images/visual associations at each locus, instead of one. This was also repeated and recommended in the homework assignments of week 2.

Linking two pieces of information at each locus during encoding

Previous experimental studies of MoL training have instructed participants to visualize one image or visual association at each locus during encoding. However, after interviewing previous participants in the World championship for mnemonists, we learned that visualizing two bits of information at each locus is probably better for facilitating subsequent recall. Encoding two (or three) bits of information at each locus was said to be the preferred choice of most competing mnemonists. The argument for this is that including two images at one locus aids retrieval of both of them, since remembering one of them is hypothesized to facilitate association with and remembering of the other (e.g. instead of imagining a chair on locus #1, and a horse on locus #2, one imagines a horse on a chair on locus #1). This approach also doubles the amount of information one can remember along a loci-route. However, in the present study we did not evaluate the effect of encoding 2 versus only 1 pieces of visual information at each locus. After the program was over, we asked a random subset of the participants whether the visualization of 2 images at each locus was perceived as problematic/difficult (yes/no; n = 8). 6 reported no difficulties, 2 reported that this was difficult.

Session 3

In the third session, the role of the hippocampus was addressed in relation to aging and MoL training: Specifically, we addressed that the hippocampal formation, and the functions that this structure is associated with, spatial memory and learning, undergoes age-related changes. The age-related changes and the probable role of this structure in MoL use, was addressed to provide rationale and motivation for performing MoL training of long word-

lists. Finally, we instructed the participants in another aspect of MoL use, namely the importance of keeping a steady rhythm in MoL encoding. When participants had a list of concrete words, or visualization of words to be encoded, they were urged to visualize them in their route at a steady and even pace. If they didn´t feel that they had encoded the images properly, they were urged to go on, and repeat encoding of the entire list, instead of stopping up at one locus and spending more time elaborating on certain images.

Session 4

In session 4, we addressed three practical memory principles, namely visualization of to-be-remembered information. Second, the use of associations to aid remembering was used. Third, the principle of organizing information, e.g. into a loci-route or as an acronym or in other ways to better remember was addressed.

Session 5

The goal of session 5 was to familiarize participants with the recently held view in neuroscience, that the brain changes throughout life in response to learning and experience, in contrast to the more traditional, static view of the brain. We illustrated how recent evidence had shed light on this matter by reporting the findings from the juggling-paper in Nature (Draganski, et al., 2004).

The second goal of session 5 was to give a brief introduction to the role of attention in memory processes. The instructor discussed facilitators (e.g. personal motivation, interest) and obstacles (stress, mood, lack of interest, mind wandering) for focused attention. The instructor told the story of the misplaced violin to illustrate the importance of attention (or lack thereof) in everyday remembering (see p. 43 of the Seven sins of memory, 2nd edition, by Schachter).

Session 6

Session 6 recapitulated on the role of attention. A simple stress-reduction technique and acronym mnemonic was introduced in class: The acronym **SOAL** (Stop, Observe, Accept, Let go) help participants to remember and utilize basic qualities of stress-reduction practice

if needed. Words like mindfulness or meditation was not used; instead the participants were simply told that it was a technique for focusing attention to "here and now".

The technique can be used before starting a task or when distracted by unwanted thoughts ("this is boring", "I can never remember this list of words", etc.) or potential distractors in the environment (irritating noise from machine outside the building, baby crying, etc.). The technique require the participants to intentionally focus attention to the present moment (**s**top, by directing attention to e.g. their breath), **o**bserve and then **a**ccept whatever thoughts, emotions or perceptions that are present, before **l**etting go and bringing attention back to e.g. the next assignment. The instructor told participants to use the technique if they felt it gave meaning to them. The technique (or the elements of the acronym) was also inserted in the beginning of the homework assignments from this week on.

Session 7

The main focus of the seventh session was the remembering of names and faces. A simple 3-step acronym and technique was used – **3A**: 1) Pay **a**ttention at the time you greet the person or hear the name for the first time, 2) Create an instant **a**ssociation with the name and/or face 3) Spatially **a**ttach – i.e., link and visualize the name (or association for the name) and person with the place you met him/her. (The acronym was 3F in Norwegian, meaning the same: **F**ølg med, **F**orbind, **F**est). The number "3" was included to remind participants of the crucial role of repetition in remembering, especially important for names. Participants were told that 3 times does it, meaning that if you repeat the name out loud when you first hear it, and use it 2 more times (either by saying it out loud when addressing the person, or repeating it for yourself), chances that you will recall it later increase substantially.

Session 8

Session 8 included a course review and teaching an acronym of the major takeaways from the training program. This acronym, KRAVSTOR, was in Norwegian, and might give little meaning for the English-speaking reader. Briefly, it is summarized as successful memory being demanding (kravstor); that no magic bullets exists, but that you can improve your memory by implementing various principles for good everyday memory – including context (**k**ontekst), you´re chance of retrieval increases if you memorize or seek out the context of

encoding (time and location). Second, you absolutely have to **r**epeat to your self what you want to remember. Repetition is a prerequisite for most successful learning. Third, **a**ssociations can be useful. Fourth, making a visual image of the stimuli (e.g. word) in your mind´s eye, increases chances of successful encoding and retrieval. Fifth, associate the stimuli you want to remember with a location (a locus; **s**ted). Sixth, excellent memory comes from **t**raining or nurture, not merely nature; You need to exercise your memory. Seventh, attention (**o**ppmerksomhet) is crucial for successful encoding in the first place; absentmindedness is the enemy of remembering. Eight, give your self praise (**r**os) and congratulate your self when you remember something right; beating your self up for having a poor memory makes it worse.

Papers I-III

I

II

III

Title: Hippocampal subfield volumes correlate with memory training benefit in subjective memory impairment

Authors: Andreas Engvig[a], Anders M. Fjell[a,b], Lars T. Westlye[a], Nina V. Skaane[c], Øyvind Sundseth[b], Kristine B. Walhovd[a,b]

Affiliations: [a]Center for the Study of Human Cognition, Department of Psychology, University of Oslo, Norway; [b]Department of Physical medicine and rehabilitation, Unit of neuropsychology, Oslo University Hospital, Ullevaal, Norway; [c]Memory clinic, Department of Geriatrics, Oslo University Hospital, Ullevaal, Norway.

Corresponding author: Andreas Engvig, Center for the Study of Human Cognition, Department of Psychology, University of Oslo, PB 1094 Blindern, 0317 Oslo, Norway.

Telephone: +47 22 84 50 58, fax: +47 22 84 50 01

e-mail address: andreas.engvig@gmail.com

Abstract

Some memory clinic outpatients still free of dementia can experience positive effects of cognitive intervention, but evidence regarding who benefit from training is lacking. Automated hippocampal volumetry might be used to foresee treatment outcomes. We hypothesized that larger hippocampal volumes are associated with greater memory performance changes following training, and that effects are selectively related to specific hippocampal subfields. 19 participants with subjective memory impairment (mean age = 60.9 years) underwent MRI-scanning and then followed an eight weeks training scheme aimed at improving verbal memory. We assessed verbal memory before and after training, and tested whether pretraining hippocampal volumes were related to memory improvements. To delineate regional specificity, we employed a new technique enabling automated volumetry of seven hippocampal subfields – including the cornu ammonis (CA) sectors and the dentate gyrus (DG). The results showed that larger hippocampal volumes before training were related to greater verbal recall improvements. Subfield volumetry revealed specific correlations between memory improvement and pretraining volumes of the left CA2/3 and CA4/DG. Depressive symptoms further gave a unique contribution in predicting gain of the intervention, independent of hippocampal volume. The results indicated that subjects with stronger depressive symptom load benefited more from training. A prediction model including baseline CA2/3-volume and depressive symptoms explained 42% of the variation in recall improvement. Our results are the first to suggest that hippocampal subfield volumetry is related to intervention outcomes in older adults experiencing memory problems. Also, previous studies have tended to exclude patients with concomitant depressive symptoms and memory complaints. The present results, however, strengthen the rationale and potential for cognitive intervention in these patients.

Keywords

Cognitive intervention, memory-training, subjective memory impairment, Hippocampal volume, hippocampal subfields, MRI volumetry

Abbreviation footnote

CA = Cornu ammonis; CVLT-II = California verbal learning test II; DG = dentate gyrus; EMQ = Everyday memory questionnaire; GDS = the 30-item Geriatric depression scale; ICV = total intracranial volume; IQ = Intelligence quotient; MCI = Mild cognitive impairment; MHI-5 = the five-item Mental health inventory from the SF-36 questionnaire; MMSE = Mini mental state examination; MP-RAGE = Magnetization prepared rapid gradient echo pulse sequence; TBV = total brain volume; WASI = Wechsler abbreviated scale of intelligence.

1.0 Introduction

Memory complaints are common in the elderly, and are associated with depressive symptoms (Reid and MacLullich, 2006), hippocampal atrophy (Saykin et al., 2006), and increased risk of developing Alzheimer´s disease (Jonker et al., 2000). For a proportion of elderly seeking help for memory complaints, the subjectively experienced deficits cannot be objectively validated – their neuropsychological profiles lie within normal range. This group of individuals straddles the boundary between normal aging and mild impairment, and is often referred to as having subjective memory impairment (Jessen et al., 2010).

A growing body of research has demonstrated positive effects of cognitive training in older adults with memory problems (Buschert et al., 2010). These studies point to cognitive intervention as a promising remediation approach for the increasing number of senior individuals seeking help for their memory. The efficiency and impact of cognitive interventions profit from evidence-based evaluations and selection of patients who might benefit most from the relevant training regimen. However, such evidence is scarce. Except for advanced age, demographical and behavioral variables have provided little or no value in predicting benefits of training (Belleville et al., 2006).

Today, reliable volumetric estimations of brain regions implicated in cognitive aging – including the hippocampus – are possible using standard MR images (Fischl et al., 2002; Hanseeuw et al., 2011). In a multi-sample study including 883 individuals, Walhovd and colleagues (2011) found consistent hippocampal volume decline in cognitively normal elderly in five of six samples (but see Sullivan et al., 2005). More progressive deterioration of the hippocampus has been shown in not only dementia (Schuff et al., 2009), but also subjective memory impairment (Striepens et al., 2010). In a cross-sectional study of 83 older adults, poorer subjective memory and better verbal recall performance on the CVLT-II were

independently associated with greater MRI-estimated thickness in the left medial temporal lobe (Bjornebekk et al., 2010). More recently, new segmentation methods haven enabled more specific delineation of the medial temporal lobe region. Hanseeuw and colleagues showed that in amnestic MCI, episodic memory was positively related to both total hippocampal volume, as well as specific subfield volumes, with the strongest effects related to CA4/DG (Hanseeuw et al., 2011). Despite the potential of MRI volumetry suggested by these studies, it is not known whether volumetry can be used to predict memory change following intervention for older adults experiencing memory problems.

The goal of the present study was to investigate associations between hippocampal volumes and memory training benefits in 19 memory clinic outpatients with subjective memory impairment. All subjects underwent MR-scanning and then followed eight weeks of memory training. The training aimed at improving verbal episodic memory, and has previously been shown to improve verbal memory and be associated with cortical thickness in cognitively healthy participants (Engvig et al., 2010). In the present study, verbal recall memory was assessed with the CVLT-II before and after the intervention.

First, we hypothesized that larger hippocampal pretraining volumes would be associated with greater verbal recall improvements. Second, since memory complaints are associated with depression (Iliffe and Pealing, 2010), we tested the impact of depressive symptom load on recall improvement. Third, hippocampus was chosen due to its well-established role in the type of memory trained. However, the hippocampal structure is anatomically and functionally heterogeneous (Yassa and Stark, 2011). In order to untangle regional specificity in the association between memory improvement and hippocampal volumes, we performed hippocampal subfield volumetry. Specifically, we calculated the volumes of CA1, CA2/3, CA4/DG, presubiculum, subiculum, fimbria and the hippocampal fissure, respectively, using a new automated procedure (Van Leemput et al., 2008, 2009). We hypothesized that effects

on memory change were likely selective to some of these regions. Cross-sectional studies have found associations between the CA2/3, CA4/DG and subicular regions and episodic memory (Hanseeuw et al., 2011; Shing et al., 2011), and we assessed whether similar relationships exist with memory change using linear regressions.

2.0 Material and methods

2.1 Participants

The study was approved by the Eastern Norway ethical committee for medical research, and informed consent was obtained from all subjects included in the study. All subjects were community-dwelling Oslo area memory clinic outpatients. Participants were referred to the clinic by their general practitioner or a specialist in neurology on the basis of self-reported memory deficits, and without knowledge of the training program. Exclusion criteria were prior history of stroke or other severe neurological or psychiatric disorders, and factors contraindicating MRI. Originally, 20 subjects were included, but one discontinued the intervention in the second week of the program. The final sample included 19 (nine females) subjects with memory complaints (lasting less than 10 years) aged 42-77 years (mean = 60.9 ±10.4 years). The age range is quite typical for this patient group, including more middle-aged adults than found in patient groups with objective cognitive deficits.

The subjects were evaluated as non-demented by the examining memory clinic physician based on ICD-10 criteria for dementia. All participants performed within the normal range on tests of general intellectual abilities as indicated by IQ > 85 as estimated from the matrices and vocabulary subtests from WASI (Wechsler, 1999). Further, all scored > 26 (mean = 29.1 ±0.9; range, 27–30) on the MMSE (Folstein et al., 1975), and ≥ 1.5 SD below the mean of age- and sex-standardized population norms on CVLT-II short and long-delay free recall (Delis et al., 2000), indicating a non-demented and relatively well-functioning study sample.

The presence of memory complaints was based on subjective reports, but confirmed in most cases by the observations of spouses or near relatives. We administered the EMQ (Sunderland et al., 1984) to provide a quantitative measure of the subjective complaint load, and compared the results with previous literature of older adults without memory concerns. According to our memory clinic register, 27% of referrals were diagnosed with subjective memory impairment in 2010 (Braekhus et al., 2011). This suggests that the present sample represents a significant proportion of the total memory clinic population.

2.2 Assessment of depressive symptoms

Depressive symptoms were assessed using the GDS (Yesavage et al., 1982) and MHI-5 (Ware and Sherbourne, 1992). Both questionnaires are validated screening tools for middle-aged and elderly populations: Using a cut-off of ≥ 14, GDS has a sensitivity and specificity for depression of 80% and 100%, respectively (Yesavage et al., 1982). For subjectively impaired older patients, cut-off for MHI-5 for depression was estimated to ≤ 47 in one study (Friedman et al., 2005). Two participants in the current sample had an MHI-5 score < 47 and GDS score > 14. Since depressive symptoms are common among patients with memory complaints (Reid and MacLullich, 2006), and these were neither taking anti-depressive medication nor reported chronic depression, they were not excluded on the basis of their depressive symptoms. Instead, we tested for the effect of depressive symptom load in the statistical analyses, also assessing whether the two participants who fell below cut-off for depression were outliers overtly influencing the results. Studentized deleted residuals from all regression analyses reported below were > - 2, and < 2 for both participants, and they were not excluded.

2.3 Intervention

An eight-week memory-training program was designed for the present study, inspired by the work of Belleville and colleagues (2006). Each week consisted of a one-hour class session and four individual homework assignments. The program focused on improving verbal recall

memory by implementing the mnemonic technique method of loci (Bower, 1970), but also taught participants everyday memory strategies and enabled group discussions. See Supplemental document 1 for a complete description.

2.4 Measurement of memory performance at baseline and follow-up

We used CVLT-II long-delay (20-minutes) free recall as a measure of verbal recall performance at both assessments. CVLT-II is a standardized neuropsychological test for various aspects of verbal learning and memory performance (Delis et al., 2000). A 16-item word list is first presented to the subject. The subject is then told to recall as many words as possible. This procedure is repeated four additional times. Following the five presentations of the initial word list, a new word list is presented once, followed by immediate free recall of that list. Next, free recall of the original word list takes place, followed by cued recall of the original word list. Following a 20-minutes delay, during which time other activities are carried out, there is a final delayed free recall trial of the original word list (long-delay recall). Alternative versions of the CVLT-II were administered at follow-up in order to minimize test-learning effects. Each participant was tested at screening for study eligibility, on average 20.4 days (SD = 15.6; range, 5–60) before training, and again the week following the completion of the training program. At the time of post-testing, participants were instructed that they could make use of the strategies learned during the intervention.

2.5 MRI acquisition and processing

MRI data were collected (8.5 ±10.7 days prior to training; range, 2–43) using a 12-channel head coil on a 1.5 T Siemens Avanto scanner (Siemens Medical Solutions, Erlangen, Germany). The pulse sequence used for morphometric analyses were two 3D T1-weighted MP-RAGE with the following parameters: TR/TE/TI/FA = 2400 ms/3.61 ms/1000 ms/8°, matrix 192 × 192, field of view = 240, 160 sagittal slices with voxel size 1.25 × 1.25 × 1.20 mm. Each scan took 7 min 42 s. Raw datasets were de-identified and transferred to Linux

workstations for processing and analyses. All scans were reviewed for quality, and automatically corrected for spatial distortion due to gradient nonlinearity (Jovicich et al., 2006) and B1 field inhomogeneity (Sled et al., 1998). The two MP-RAGE volumes were averaged to increase signal-to-noise ratio and brain volume estimation reliability.

2.6 Volumetric analysis

All brain volumes were estimated using FreeSurfer (http://surfer.nmr.mgh.harvard.edu/). First, the whole hippocampal formation was segmented using the standard segmentation procedure (Fischl et al., 2002; Fischl et al., 2004). The procedure automatically labels each voxel in the brain as one of 40 structures (Fischl et al., 2002) using a probabilistic brain atlas specific for the current image acquisition protocol (Han et al. 2006). In older subjects scanned on the same Siemens MR scanner, FreeSurfer is shown to calculate consistent hippocampal formation volumes from 1.5 T MP-RAGE images with reproducibility errors of 3.6% and 3.4% for the left and right hippocampi, respectively (Jovicich et al., 2009). Next, we performed automated subfield segmentation of the hippocampus using a new technique within the FreeSurfer suite. The procedure uses Bayesian inference and a probabilistic atlas of the hippocampal formation based on manual delineations of subfields in ultra-high T1-weighted MRI scans from a number of different subjects (Van Leemput et al., 2009). Seven subfield volumes are calculated: CA1, CA2/3, CA4/DG, presubiculum, subiculum, hippocampal fissure, and fimbria. The larger subfields are shown to correlate well with manual volume estimates, with an average dice coefficient of around 0.74 for CA2/3 and subiculum (Van Leemput et al., 2009). Segmentation results were visually inspected for errors in all datasets, and no manual edits were done. Please see Figure 1b for whole hippocampus and subfields segmentation results in one of the participants.

Volumes of the whole hippocampus, as well as hippocampal subfield volumes, were adjusted for TBV and the residuals used in the statistical analyses. TBV was defined as the sum of the

volumes of cerebral and cerebellar grey and white matter, excluding the ventricles, cerebrospinal fluid and dura which are included in ICV. TBV and ICV are highly correlated, and often used to account for sex differences in global brain or head size. However, TBV is deemed more sensitive to global brain atrophy, due to e.g. age or disease, and was used here in order to delineate specific anatomical associations beyond global effects. For scatterplots of individual data (Figure 1a), we calculated normative measures of the regional volumes. The following formula was used (Jack et al., 1989): $Volume_{adjusted} = Volume_{observed} - \beta$ [slope from TBV vs. regional volume regression] * ($TBV_{observed} - TBV_{sample\ mean}$).

2.7 Statistical analyses

PASW Statistics 18.0 (SPSS Inc., Chicago, Ill) was used for the statistical analyses. First, CVLT-II long-delay recall performance changes were tested by means of paired samples t-tests. Next, we investigated relationships between the neuroanatomical pretraining volumes and CVLT-II long-delay recall change. Recall change was calculated as the difference in recall performance between baseline and follow-up. All pretraining volumes, long-delay recall change, and age and GDS (covariates, see below) were deemed normally distributed (Shapiro-Wilk´s tests; p-values > 0.050). Long-delay recall change scores were adjusted for baseline performance by means of linear regressions, and the standardized residuals were used in the analyses. We tested whether hippocampal volume was related to recall performance change, using multiple linear regressions. Change residuals for CVLT-II long-delay free recall was included as a dependent variable. In the first model, hippocampal volume was included as a predictor variable. We performed separate analyses for each hemisphere to assess laterality effects. In a second model, we included GDS as an additional predictor to test whether depressive symptoms co-varied with recall change. Finally, in follow-up analyses, we included each of seven subfield volumes to delineate subfield specificity. Age and sex was included as covariates in all models.

3.0 Results

Participation rate was high. On average, participants completed 96.7% of the 40 in-class and homework exercises (SD = 5.1). Mean EMQ-score was 105.4 (SD = 38.2). In comparison, mean EMQ-score in a sample of 83 healthy older adults without memory concerns was 63.1 (SD = 24.3) (Bjornebekk et al., 2010), indicating increased subjective memory concerns in the current sample. Mean (SD) baseline and re-test performance on CVLT-II long-delay free recall were 11.3 (3.1) and 14.1 (2.6), respectively. Paired samples t-tests confirmed a significant mean increase in long-delay free recall ($df = 18$, $t = 5.0$, $p < 0.001$). Four individuals had a long-delay recall change score ≤ 0, indicating that these participants had no benefit of training and re-test. Mean (SD) GDS and MHI-5 score were 9.2 (6.2) and 70.3 (19.4), respectively.

3.1 Larger left hippocampal volume and more depressive symptoms are independently associated with greater recall change

Results from the multiple linear regressions predicting free verbal recall change from hippocampal and subfield volumes, respectively, are presented in Table 1. Figure 1a shows scatterplots of significant associations.

Separate regression analyses with left and right hippocampal volume included as a predictor revealed a selective effect of the left hippocampus on recall change ($\beta = 0.73$, $t = 2.64$, $p = .019$). The result indicates that larger baseline left hippocampal volume was related to more positive free recall change. A second model with left hippocampal volume and GDS on recall change further revealed a significant effect of depressive symptoms ($\beta = 0.41$, $t = 2.18$, $p = .047$), indicating that stronger depressive symptomatology at baseline was related to more positive free recall changes. Importantly, both left hippocampal volume and GDS provided unique statistical contribution to this regression model. The full model explained almost 38%

of the variation in recall change (F = 3.75, p = .028, adjusted R^2 = .379). We found no association between GDS and left hippocampal volume (Pearson´s *r* = -.08, p = .74).

3.2 Effects on recall change are selectively related to left CA2/3 and CA4/DG subfields

As effects were selective to the left hemisphere, we proceeded with follow-up analyses of left subfield volumes. Two regions showed significant effects on recall change, namely the left CA2/3 and CA4/DG (see Table 1). GDS gave trend contributions to the regression models (p-values < 0.06). The models including GDS explained 42% (left CA2/3 model; F = 4.28, p < 0.02, adjusted R^2 = .421) and 39% (left CA4/DG model; F = 3.88, p < 0.03, adjusted R^2 = .391) of the variation in free recall change, respectively.

None of the other subfields were related to recall change (p-values > 0.10). Further, there were no associations between GDS and CA2/3 (Pearson´s r = -.09, p = .72) and CA4/DG (Pearson´s *r* = -.02, p = .94) volume, respectively. We found no unique contributions of age or sex in any of the tested models of free recall change (p-values > 0.10).

4.0 Discussion

We report novel results demonstrating that hippocampal volume and depressive symptoms are both related to memory changes following cognitive intervention in outpatients with subjective memory impairment. Hippocampus is known to be heavily involved in episodic memory (Tulving and Markowitsch, 1998), but here we show for the first time that baseline volumes of the hippocampus, including select subfields, can be used to predict how much elderly patients with subjective memory impairment benefit from memory training. The significance of depressive symptomatology points to the importance of emotional status in evaluating potential for intervention. Implications of the findings are discussed.

4.1 General implications

The use of MRI-derived medial temporal lobe volumes has previously been shown valuable in diagnostic predictions of conversion from MCI to Alzheimer´s disease (Jack et al., 1999), as well as in prediction of memory decline in healthy elderly (Murphy et al., 2010). The present findings suggest that hippocampal volumetry might also serve to foresee intervention outcomes in a large clinical group for which few treatment options presently exist. This patient group is also interesting because it for some of the patients represents an early transitional phase between normal functioning and cognitive impairment. Those who eventually develop MCI or dementia are, when experiencing subjective memory impairment, possibly at a stage of the disease so early that if effective treatment was available, adverse effects could potentially be halted, minimizing subsequent neurodegeneration and cognitive decrements. Even at the stage of amnestic MCI, brain degeneration is so advanced that it is less conceivable that treatment will fully reverse atrophy and restore cognitive function to a preclinical level. For otherwise healthy elderly with memory complaints, the window of intervention is likely not closed, and more knowledge about factors influencing the potential for memory improvement is thus vital.

4.2 Hippocampal volumetry and verbal recall changes

The results demonstrate that larger hippocampal volumes are related to greater increases in recall performance after cognitive intervention for subjective memory impairment. Previous cross-sectional (Walhovd et al., 2004) and longitudinal (Murphy et al., 2010) studies of healthy elderly have also documented positive relationships between CVLT recall scores and hippocampal volume (but see Van Petten et al., 2004).

The current results were selective to the left hemisphere. According to one study, the left hippocampal region degenerates before and faster in Alzheimer´s disease compared to the right (Thompson et al., 2003). If neurodegenerative atrophy indeed develops earlier in the left hemisphere, this could possibly affect some of the subjects in the present study whose

subjective memory complaints represent pre-clinical dementia, in turn mediating variation in recall change. Longitudinal diagnostic follow-up of participants is necessary in order to shed light on this hypothesis. Nonetheless, the selectiveness of the left hippocampus is also supported by the left-hemisphere dominance for language (Gazzaniga, 1995) and previous cross-sectional reports of verbal recall and volume (Ystad et al., 2009).

In the present study, we were able to delineate training-related hippocampal effects to distinct subfields. Subfield volumetry suggested that the effects on verbal recall change were selective for the left CA2/3 and CA4/DG. This adds to recent cross-sectional subfield volumetry studies of these regions in episodic memory (Hanseeuw et al., 2011; Shing et al., 2011). The finding further strengthens the present conclusions, as both CA3 and DG are functionally highly relevant to the trained task. Both regions play key roles in pattern separation (Bakker et al., 2008), a hallmark feature of episodic memory function (Yassa and Stark, 2011).

Next, there are at least two plausible explanations for the finding of a positive relationship between change in memory function and hippocampal volumes: First, hippocampal volume has been shown to be the brain measure that distinguishes best between patients with MCI and healthy elderly (Fjell et al., 2010), but is also reduced in healthy aging (Fjell et al., 2009; Walhovd et al., 2011). Including the present subfields volumetry technique further increases the discriminative power in detecting subjects with MCI (Hanseeuw et al., 2011). Atrophic processes related to normal aging or degenerative disease in its earliest stages will to some extent be reflected in these volumes. Patients with the greater hippocampal volume are likely to have larger degree of neuronal and cognitive plasticity, and thereby more potential for restoration and improvement of cognitive functions in response to targeted cognitive intervention. This hypothesis is corroborated by the fact that we found associations with recall change and the DG, one of the very few brain areas where adult neurogenesis has been documented (Ming and Song, 2011).

A second explanation for the positive volume / recall change relationships is that hippocampal size may represent a reserve factor for cognitive function. Thus, greater hippocampal volume is not necessarily beneficial for memory improvements because large volume is related to lower rates of ongoing atrophy or increased plasticity, but rather because a large initial hippocampus makes the person better able to sustain negative impact from aging or early degenerative processes. The two hypotheses are not mutually exclusive, and longitudinal studies spanning longer time intervals are necessary to disentangle the relative contributions of these alternative explanations.

4.3 Relationship with depressive symptom load

Although the present intervention was a cognitive training program, several aspects of the program can be envisioned to have positive affective side effects, e.g. repeated social gatherings and interactions with the experimenter. Previous studies have shown that depressive symptoms have adverse effects on recall performance (Kizilbash et al., 2002), and increase of depressive symptom load is often seen in elderly accompanying memory problems of uncertain origin (Reid and MacLullich, 2006). We found no correlation between depressive symptoms and baseline recall performance in the present study (r = .13, p = .60). We did, however, find that stronger baseline depressive symptoms were associated with stronger recall improvement. Larger hippocampal volumes and more depressive symptoms predicted recall improvement independently of each other, and we found no correlation between depression symptom load and hippocampal volume. Also, none of the participants reported chronic depression shown to reduce hippocampal volume (MacQueen et al., 2003). A possible interpretation of our finding is that group-based intervention relieves memory constraints due to depressive symptomatology. Subjects with no or few depressive symptoms, but whose subjective memory impairment rather reflects incipient neurodegenerative disease, might benefit less from training due to pathological neuronal constraints. As the study was designed

to assess correlates between hippocampal volume and memory change, we did not assess depressive symptoms longitudinally. We regret this, and further investigation is needed in order to delineate the relationship between affective health and training benefit.

4.4 Methodological considerations and perspectives for future trials

The present study did not include an active control condition to test the efficacy of the program. Since the present effects are identified within a group undergoing training, the main finding that hippocampus is related to change in memory performance within this group is nonetheless valid. But we cannot assess to what extent other factors, such as practice effects drives the observed recall change. However, several reasons suggest that the recall change is more than a mere effect of taking the test twice. First, significant verbal memory gains following the present intervention is documented in a randomized controlled study of healthy elderly (Engvig et al., 2010). Second, we opted to reduce practice effects by using alternate versions of the memory outcome measure. Counterbalancing alternate and original forms between baseline and follow-up could further reduce variability due to difference in forms, and should be implemented in future trials. Third, a similar eight-week memory training protocol with memory clinic attendees free of dementia showed a 5-14% decrease in recall at retest in a wait-list condition (Belleville 2006). In the present study we observed a mean increase in recall, despite four subjects having no improvement. Also, a number of previous controlled studies that have documented effects of cognitive intervention such as the present (for a review, see Buschert et al., 2010).

The present study included short-term assessments only. Other investigators have found that training effects in older adults can last up to five years (Willis et al., 2006), but how individual differences in affective health and brain volumetry impacts the stability of these effects is not known. An exciting future enquiry is whether the present cognitive intervention has an effect on the depressive symptoms shown to be associated with older adults´ memory

concerns (Iliffe and Pealing, 2010). MRI studies of individuals with subjective memory complaints have tended to exclude those with concomitant depression and memory concerns (Saykin et al., 2006). In our view, when studying remediation approaches for older adults with memory complaints, subjects with mild-moderate depressive symptoms should not always be excluded. In fact, several studies have shown that effects of cognitive intervention are partly mediated by alleviating affective symptom load (Rozzini et al., 2007; Talassi et al., 2007). As depression has been associated with more than twofold increase in dementia risk (Byers and Yaffe, 2011), our findings should not reduce, but instead further strengthen the rationale for cognitive intervention in patients with memory complaints and depressive symptoms.

A strength of the present study is the high participation rates, indicating that the present intervention was feasible. However, the high general cognitive level with mean IQ > 1 SD above the estimated population mean, implies that the study cannot address whether samples with lower IQ would benefit from interventions such as the present, or whether the current prediction models could be used for patients with lower general functioning. Future intervention studies should aim to recruit samples with broader ranges of cognitive functioning.

The final considerations relate to the current MRI-segmentation procedures. In the present study, the mean estimated beta-values were larger for the whole hippocampal volume prediction model, as compared to the CA2/3 and CA4/DG models (Table 1). Still, although not formally tested statistically, better apparent overall model fits were obtained for the two subfield prediction models as compared to the model including the whole left hippocampus. In the validation study of this technique, the CA2/3 and CA4/DG showed the highest reliability (Van Leemput et al., 2009). It is possible that different reliability between the subfield measures can contribute to the stronger memory change correlations seen for CA2/3

and CA4/DG. In the present study, we used 1.5 T scans as compared to the 3 T images used for the development of the subfield technique employed. It is unknown which effect this difference in field strength has on the current segmentation results. Visual inspection of our results (for an example, see Figure 1b) suggests subfield identification and separation in agreement with results reported at 3 T (see Fig. 2, p. 3 in Hanseeuw et al., 2011). Thus, the current FreeSurfer algorithm seems to provide adequate segmentation results even at 1.5 T. Nevertheless, across field-strength validation studies are surely awaited.

4.5 Conclusion

We present novel findings showing that training-related changes in memory performance are related to left CA2/3 and CA4/DG size in outpatients with subjective memory impairment. Hippocampal volumes and depressive symptom load were independently related to individual differences in verbal recall changes. The results support a selective relationship between memory training benefit and hippocampal size in older adults experiencing memory problems. The findings also points to the usefulness of MRI-volumetry in predicting relevant intervention outcomes, previously supported by studies of healthy populations (e.g. Erickson et al., 2010).

5.0 Acknowledgements

The authors acknowledge the Research Council of Norway for an operating grant to AMF and KBW (Grant number: 186092), and the Kavli Research Center for Aging and Dementia for salary support to AE. We thank the staff at the Memory clinics, Oslo University hospital and Bærum hospital for aiding patient recruitment, Dr. Anne Rita Øksengaard for diagnostic evaluations and patient recruitment, and Dr. Paulina Due-Tønnessen for neuroradiological evaluations of all participants.

6.0 Disclosure Statement

All authors declare no actual or potential conflicts of interest.

References

Bakker, A., Kirwan, C.B., Miller, M., Stark, C.E., 2008. Pattern separation in the human hippocampal CA3 and dentate gyrus. Science 319, 1640-1642.

Belleville, S., Gilbert, B., Fontaine, F., Gagnon, L., Menard, E., Gauthier, S., 2006. Improvement of episodic memory in persons with mild cognitive impairment and healthy older adults: evidence from a cognitive intervention program. Dementia and geriatric cognitive disorders 22, 486-499.

Bjornebekk, A., Westlye, L.T., Walhovd, K.B., Fjell, A.M., 2010. Everyday memory: self-perception and structural brain correlates in a healthy elderly population. Journal of the International Neuropsychological Society : JINS 16, 1115-1126.

Bower, G.H., 1970. Analysis of a mnemonic device. American Scientist 58, 496-510.

Braekhus, A., Ulstein, I., Wyller, T.B., Engedal, K., 2011. The Memory Clinic - outpatient assessment when dementia is suspected. Tidsskrift for den Norske laegeforening 131, 2254-2257.

Buschert, V., Bokde, A.L., Hampel, H., 2010. Cognitive intervention in Alzheimer disease. Nature reviews. Neurology 6, 508-517.

Byers, A.L., Yaffe, K., 2011. Depression and risk of developing dementia. Nature reviews. Neurology.

Delis, D., Kramer, J., Kaplan, E., Ober, B., 2000. California Verbal Learning Test-Second Edition (CVLT-II) Manual. San Antonio: The Psychological Corporation.

Engvig, A., Fjell, A.M., Westlye, L.T., Moberget, T., Sundseth, Ø., Larsen, V.A., Walhovd, K.B., 2010. Effects of memory training on cortical thickness in the elderly. NeuroImage 52, 1667-1676.

Erickson, K.I., Boot, W.R., Basak, C., Neider, M.B., Prakash, R.S., Voss, M.W., Graybiel, A.M., Simons, D.J., Fabiani, M., Gratton, G., Kramer, A.F., 2010. Striatal volume predicts level of video game skill acquisition. Cerebral cortex 20, 2522-2530.

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341-355.

Fischl, B., Salat, D.H., van der Kouwe, A.J., Makris, N., Segonne, F., Quinn, B.T., Dale, A.M., 2004. Sequence-independent segmentation of magnetic resonance images. NeuroImage 23 Suppl 1, S69-84.

Fjell, A.M., Walhovd, K.B., Fennema-Notestine, C., McEvoy, L.K., Hagler, D.J., Holland, D., Brewer, J.B., Dale, A.M., 2009. One-Year Brain Atrophy Evident in Healthy Aging. J. Neurosci. 29, 15223-15231.

Fjell, A.M., Walhovd, K.B., Fennema-Notestine, C., McEvoy, L.K., Hagler, D.J., Holland, D., Brewer, J.B., Dale, A.M., 2010. CSF biomarkers in prediction of cerebral and clinical change in mild cognitive impairment and Alzheimer's disease. The Journal of neuroscience : the official journal of the Society for Neuroscience 30, 2088-2101.

Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. Journal of psychiatric research 12, 189-198.

Friedman, B., Heisel, M., Delavan, R., 2005. Validity of the SF-36 five-item Mental Health Index for major depression in functionally impaired, community-dwelling elderly patients. Journal of the American Geriatrics Society 53, 1978-1985.

Gazzaniga, M.S., 1995. Principles of human brain organization derived from split-brain studies. Neuron 14, 217-228.

Hanseeuw, B.J., Van Leemput, K., Kavec, M., Grandin, C., Seron, X., Ivanoiu, A., 2011. Mild cognitive impairment: differential atrophy in the hippocampal subfields. AJNR. American journal of neuroradiology 32, 1658-1661.

Iliffe, S., Pealing, L., 2010. Subjective memory problems. BMJ 340, c1425.

Jack, C.R., Jr., Petersen, R.C., Xu, Y.C., O'Brien, P.C., Smith, G.E., Ivnik, R.J., Boeve, B.F., Waring, S.C., Tangalos, E.G., Kokmen, E., 1999. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. Neurology 52, 1397-1403.

Jack, C.R., Jr., Twomey, C.K., Zinsmeister, A.R., Sharbrough, F.W., Petersen, R.C., Cascino, G.D., 1989. Anterior temporal lobes and hippocampal formations: normative volumetric measurements from MR images in young adults. Radiology 172, 549-554.

Jessen, F., Wiese, B., Bachmann, C., Eifflaender-Gorfer, S., Haller, F., Kolsch, H., Luck, T., Mosch, E., van den Bussche, H., Wagner, M., Wollny, A., Zimmermann, T., Pentzek, M., Riedel-Heller, S.G., Romberg, H.P., Weyerer, S., Kaduszkiewicz, H., Maier, W., Bickel, H., 2010. Prediction of dementia by subjective memory impairment: effects of severity and temporal association with cognitive impairment. Archives of general psychiatry 67, 414-422.

Jonker, C., Geerlings, M.I., Schmand, B., 2000. Are memory complaints predictive for dementia? A review of clinical and population-based studies. International Journal of Geriatric Psychiatry 15, 983-991.

Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., Kennedy, D., Schmitt, F., Brown, G., MacFall, J., Fischl, B., Dale, A., 2006. Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. NeuroImage 30, 436-443.

Jovicich, J., Czanner, S., Han, X., Salat, D., van der Kouwe, A., Quinn, B., Pacheco, J., Albert, M., Killiany, R., Blacker, D., Maguire, P., Rosas, D., Makris, N., Gollub, R., Dale, A., Dickerson, B.C., Fischl, B., 2009. MRI-derived measurements of human subcortical,

ventricular and intracranial brain volumes: Reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. NeuroImage 46, 177-192.

Kizilbash, A.H., Vanderploeg, R.D., Curtiss, G., 2002. The effects of depression and anxiety on memory performance. Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists 17, 57-67.

MacQueen, G.M., Campbell, S., McEwen, B.S., Macdonald, K., Amano, S., Joffe, R.T., Nahmias, C., Young, L.T., 2003. Course of illness, hippocampal function, and hippocampal volume in major depression. Proceedings of the National Academy of Sciences of the United States of America 100, 1387-1392.

Ming, G.L., Song, H., 2011. Adult neurogenesis in the mammalian brain: significant answers and significant questions. Neuron 70, 687-702.

Murphy, E.A., Holland, D., Donohue, M., McEvoy, L.K., Hagler Jr, D.J., Dale, A.M., Brewer, J.B., 2010. Six-month atrophy in MTL structures is associated with subsequent memory decline in elderly controls. NeuroImage 53, 1310-1317.

Reid, L., MacLullich, A., 2006. Subjective Memory Complaints and Cognitive Impairment in Older People. Dementia and geriatric cognitive disorders 22, 471-485.

Rozzini, L., Costardi, D., Chilovi, B.V., Franzoni, S., Trabucchi, M., Padovani, A., 2007. Efficacy of cognitive rehabilitation in patients with mild cognitive impairment treated with cholinesterase inhibitors. International Journal of Geriatric Psychiatry 22, 356-360.

Saykin, A.J., Wishart, H.A., Rabin, L.A., Santulli, R.B., Flashman, L.A., West, J.D., McHugh, T.L., Mamourian, A.C., 2006. Older adults with cognitive complaints show brain atrophy similar to that of amnestic MCI. Neurology 67, 834-842.

Schuff, N., Woerner, N., Boreta, L., Kornfield, T., Shaw, L.M., Trojanowski, J.Q., Thompson, P.M., Jack, C.R., Jr., Weiner, M.W., 2009. MRI of hippocampal volume loss in

early Alzheimer's disease in relation to ApoE genotype and biomarkers. Brain : a journal of neurology 132, 1067-1077.

Shing, Y.L., Rodrigue, K.M., Kennedy, K.M., Fandakova, Y., Bodammer, N., Werkle-Bergner, M., Lindenberger, U., Raz, N., 2011. Hippocampal subfield volumes: age, vascular risk, and correlation with associative memory. Frontiers in aging neuroscience 3, 2.

Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. Medical Imaging, IEEE Transactions on 17, 87-97.

Striepens, N., Scheef, L., Wind, A., Popp, J., Spottke, A., Cooper-Mahkorn, D., Suliman, H., Wagner, M., Schild, H.H., Jessen, F., 2010. Volume loss of the medial temporal lobe structures in subjective memory impairment. Dementia and geriatric cognitive disorders 29, 75-81.

Sullivan, E.V., Marsh, L., Pfefferbaum, A., 2005. Preservation of hippocampal volume throughout adulthood in healthy men and women. Neurobiology of aging 26, 1093-1098.

Sunderland, A., Harris, J.E., Gleave, J., 1984. Memory failures in everyday life following severe head injury. Journal of clinical neuropsychology 6, 127-142.

Talassi, E., Guerreschi, M., Feriani, M., Fedi, V., Bianchetti, A., Trabucchi, M., 2007. Effectiveness of a cognitive rehabilitation program in mild dementia (MD) and mild cognitive impairment (MCI): a case control study. Archives of gerontology and geriatrics 44 Suppl 1, 391-399.

Thompson, P.M., Hayashi, K.M., de Zubicaray, G., Janke, A.L., Rose, S.E., Semple, J., Herman, D., Hong, M.S., Dittmer, S.S., Doddrell, D.M., Toga, A.W., 2003. Dynamics of gray matter loss in Alzheimer's disease. The Journal of neuroscience : the official journal of the Society for Neuroscience 23, 994-1005.

Tulving, E., Markowitsch, H.J., 1998. Episodic and declarative memory: role of the hippocampus. Hippocampus 8, 198-204.

Van Leemput, K., Bakkour, A., Benner, T., Wiggins, G., Wald, L.L., Augustinack, J., Dickerson, B.C., Golland, P., Fischl, B., 2008. Model-based segmentation of hippocampal subfields in ultra-high resolution in vivo MRI. Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention 11, 235-243.

Van Leemput, K., Bakkour, A., Benner, T., Wiggins, G., Wald, L.L., Augustinack, J., Dickerson, B.C., Golland, P., Fischl, B., 2009. Automated segmentation of hippocampal subfields from ultra-high resolution in vivo MRI. Hippocampus 19, 549-557.

Van Petten, C., Plante, E., Davidson, P.S., Kuo, T.Y., Bajuscak, L., Glisky, E.L., 2004. Memory and executive function in older adults: relationships with temporal and prefrontal gray matter volumes and white matter hyperintensities. Neuropsychologia 42, 1313-1335.

Walhovd, K.B., Fjell, A.M., Reinvang, I., Lundervold, A., Fischl, B., Quinn, B.T., Dale, A.M., 2004. Size does matter in the long run: hippocampal and cortical volume predict recall across weeks. Neurology 63, 1193-1197.

Walhovd, K.B., Westlye, L.T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., Agartz, I., Salat, D.H., Greve, D.N., Fischl, B., Dale, A.M., Fjell, A.M., 2011. Consistent neuroanatomical age-related volume differences across multiple samples. Neurobiology of aging 32, 916-932.

Ware, J.E., Jr., Sherbourne, C.D., 1992. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Medical care 30, 473-483.

Wechsler, D., 1999. Wechsler Abbreviated Scale of Intelligence. The Psychological Corporation, San Antonio, TX.

Willis, S.L., Tennstedt, S.L., Marsiske, M., Ball, K., Elias, J., Koepke, K.M., Morris, J.N., Rebok, G.W., Unverzagt, F.W., Stoddard, A.M., Wright, E., for the ACTIVE Study Group, 2006. Long-term Effects of Cognitive Training on Everyday Functional Outcomes in Older Adults. JAMA 296, 2805-2814.

Yassa, M.A., Stark, C.E., 2011. Pattern separation in the hippocampus. Trends in neurosciences 34, 515-525.

Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M., Leirer, V.O., 1982. Development and validation of a geriatric depression screening scale: a preliminary report. Journal of psychiatric research 17, 37-49.

Ystad, M.A., Lundervold, A.J., Wehling, E., Espeseth, T., Rootwelt, H., Westlye, L.T., Andersson, M., Adolfsdottir, S., Geitung, J.T., Fjell, A.M., Reinvang, I., Lundervold, A., 2009. Hippocampal volumes are important predictors for memory function in elderly women. BMC Med Imaging 9, 17.

Figure 1 caption

Title: Volume/recall change relationships and hippocampal segmentation results

A) The figure shows scatterplots and linear fit-lines of three significant volume – long delay recall change relationships. All volumes are from the left hemisphere. Variability due to age, sex, depressive symptoms and baseline score has been partialled out using linear regressions. The normative volumes are shown in mm^3.

B) The figure shows the results of the hippocampus segmentation for one subject superimposed on the subject´s T1-weighted scan in axial, coronal and sagittal views, respectively. Top row: Hippocampal subfield segmentation. The hippocampal fissure is not visible, and thus not marked. Presub = Presubiculum. Sub = Subiculum. Lower row: Whole hippocampal formation segmentation.

Table 1

The table shows results from multiple linear regressions testing different prediction models of baseline-adjusted changes in CVLT-II long-delay free recall score as dependent variable.

		Volume		GDS	
Models	Adj-R^2	Beta	p	Beta	p
Whole hippocampus					
Left hemisphere	.22	**0.73**	**.019**	-	-
	.38	**0.76**	**.009**	**0.41**	**.047**
Right hemisphere	- .05	0.34	.263	-	-
	.05	0.31	.294	0.36	.144
Hippocampal subfields					
Left CA2/3	.29	**.66**	**.009**	-	-
	.42	**.66**	**.005**	.38	.054
Left CA4/DG	.26	**.67**	**.013**	-	-
	.39	**.67**	**.007**	.38	.058

Table footnote: Two separate regression models are shown for each brain region; with and without GDS (depressive symptoms) as an additional predictor variable. Long-

delay free recall is the change in this CVLT-metric, adjusted for pre-training score. Adj-R^2 denotes the adjusted full model fit, and the betas are the unique standardized regression slopes for each predictor variable. Only significant subfield models are shown. Bold indicates significance at $p < 0.05$-level. Age and sex are included as covariates in all models.

Figure 1
Click here to download high resolution image