NeuroImage 61 (2012) 188–194

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

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

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

[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

ARTICLE INFO

Article history:
Accepted 25 February 2012
Available online 4 March 2012

Keywords:
Cognitive intervention
Memory-training
Subjective memory impairment
Hippocampal volume
Hippocampal subfields
MRI volumetry

ABSTRACT

Although some older adults experiencing memory problems have been shown to benefit from cognitive training, evidence regarding who will improve from this type of intervention 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 memory clinic outpatients with subjective memory impairment (mean age = 60.9 years) underwent MRI-scanning and then followed an eight week 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 a stronger depressive symptom load benefited more from the 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.

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

Abbreviations: 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.

* Corresponding author at: Center for the Study of Human Cognition, Department of Psychology, University of Oslo, PB 1094 Blindern, 0317 Oslo, Norway. Fax: +47 22 84 50 01.
E-mail address: andreas.engvig@gmail.com (A. Engvig).

1053-8119/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2012.02.072

standard MR images (Fischl et al., 2002; Hanseeuw et al., 2011). In a multi-sample study including 883 individuals, Walhovd et al. (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 (Bjørnebekk et al., 2010). More recently, new segmentation methods have 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.

Material and methods

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.

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.

Intervention

An eight-week memory-training program was designed for the present study, inspired by the work of Belleville et al. (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.

Measurement of memory performance at baseline and follow-up

We used CVLT-II long-delay (20-minute) 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-minute 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.

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.

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 Fig. 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 gray 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 (Fig. 1a), we calculated normative measures of the regional volumes. The following formula was used (Jack et al., 1989): $\text{Volume}_{\text{adjusted}} = \text{Volume}_{\text{observed}} - \beta\ [\text{slope from TBV vs. regional volume regression}] * (\text{TBV}_{\text{observed}} - \text{TBV}_{\text{sample mean}})$.

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 were 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, we proceeded with follow-up analyses of the seven subfield-volumes to delineate subfield specificity. In order to reduce the number of statistical tests and minimize the possibility of Type I errors, we investigated subfield specificity in the given hemisphere only if the first model was deemed significant. Age and sex were included as covariates in all models.

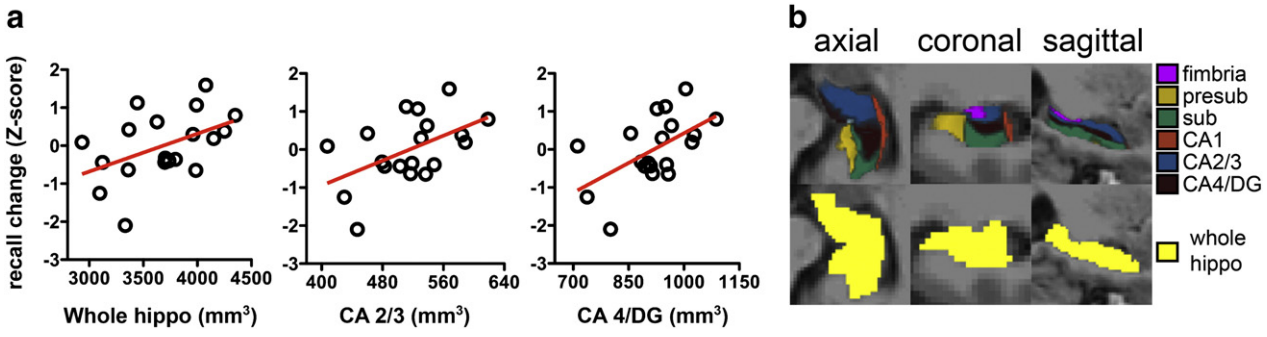

Fig. 1. Volume/recall change relationships and hippocampal segmentation results. a) The figure shows scatterplots and linear fit-lines of three significant volume–long delay recall change relationships. All volumes are from the left hemisphere. Variability due to age, sex, depressive symptoms and baseline 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.

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) (Bjørnebekk 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.

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. Fig. 1a shows scatterplots of significant associations. Note that the results represent uncorrected statistical values and have not been adjusted for multiple comparisons.

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 $r = -.08$, $p = .74$).

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).

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.

Models	Adj-R^2	Volume		GDS	
		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

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.

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.

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 for some of the patients it 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.

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.

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 an 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 a 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.

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 drive 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 are 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 et al., 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 impact 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 subfields showed the

highest reliability (Van Leemput et al., 2009). It is possible that the increased reliability of these subfields has contributed to the stronger memory change correlations seen for CA2/3 and CA4/DG. In the present study, we used 1.5 T scans (1.25×1.25×1.20 mm resolution) as compared with the 3 T images (380 μm in-plane resolution; slice thickness 0.8 mm) used for the development of the subfield technique employed. It is unknown which effect differences in field strength and image resolution have on the current segmentation results. Visual inspection of our results (for an example, see Fig. 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 and image-resolution validation studies are surely awaited.

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 point to the usefulness of MRI-volumetry in predicting relevant intervention outcomes, previously supported by studies of healthy populations (e.g., Erickson et al., 2010).

Disclosure statement

All authors declare no actual or potential conflicts of interest.

Acknowledgments

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.neuroimage.2012.02.072.

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. Dement. Geriatr. Cogn. Disord. 22, 486–499.
Bjørnebekk, A., Westlye, L.T., Walhovd, K.B., Fjell, A.M., 2010. Everyday memory: self-perception and structural brain correlates in a healthy elderly population. J. Int. Neuropsychol. Soc. 16, 1115–1126.
Bower, G.H., 1970. Analysis of a mnemonic device. Am. Sci. 58, 496–510.
Braekhus, A., Ulstein, I., Wyller, T.B., Engedal, K., 2011. The Memory Clinic — outpatient assessment when dementia is suspected. Tidsskr. Nor. laegeforen. 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 7, 323–331.
Delis, D., Kramer, J., Kaplan, E., Ober, B., 2000. California Verbal Learning Test-Second Edition (CVLT-II) Manual. The Psychological Corporation, San Antonio.
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. Cereb. 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–S84.
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. J. Neurosci. 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. J. Psychiatr. Res. 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. J. Am. Geriatr. Soc. 53, 1978–1985.
Gazzaniga, M.S., 1995. Principles of human brain organization derived from split-brain studies. Neuron 14, 217–228.
Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., Busa, E., Pacheco, J., Albert, M., Killiany, R., Maguire, P., Rosas, D., Makris, N., Dale, A., Dickerson, B., Fischl, B., 2006. Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. NeuroImage 32, 180–194.
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 Am. J. Neuroradiol. 32, 1658–1661.
Iliffe, S., Pealing, L., 2010. Subjective memory problems. BMJ 340, c1425.
Jack Jr., C.R., 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.
Jack Jr., C.R., 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.
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. Arch. Gen. 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. Int. J. Geriatr. 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. Arch. Clin. Neuropsychol. 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. Proc. Natl. Acad. Sci. U. S. A. 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. Dement. Geriatr. Cogn. Disord. 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. Int. J. Geriatr. 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 Jr., C.R., Weiner, M.W., 2009. MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. Brain 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. Front. Aging Neurosci. 3, 2.

Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. Med. Imaging, IEEE Trans. 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. Dement. Geriatr. Cogn. Disord. 29, 75–81.

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

Sunderland, A., Harris, J.E., Gleave, J., 1984. Memory failures in everyday life following severe head injury. J. Clin. Neuropsychol. 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. Arch. Gerontol. Geriatr. 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. J. Neurosci. 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, pp. 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. Neurobiol. Aging 32, 916–932.

Ware Jr., J.E., Sherbourne, C.D., 1992. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med. 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 Neurosci. 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. J. Psychiatr. Res. 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.