



Effects of memory training on cortical thickness in the elderly

Andreas Engvig^{a,b,*}, Anders M. Fjell^{b,c}, Lars T. Westlye^b, Torgeir Moberget^d, Øyvind Sundseth^c, Vivi Agnete Larsen^b, Kristine B. Walhovd^{b,c,*}

^a Medical Student Research Program, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway

^b Center for the Study of Human Cognition, Department of Psychology, University of Oslo, Oslo, Norway

^c Department of Neuropsychology, Oslo University Hospital, Ullevaal, Norway

^d Department of Neuropsychiatry and Psychosomatic Medicine, Oslo University Hospital, Rikshospitalet, Norway

ARTICLE INFO

Article history:

Received 27 January 2010

Revised 10 May 2010

Accepted 17 May 2010

Available online 24 May 2010

Keywords:

Brain
Plasticity
Aging
Imaging
Memory
Training

ABSTRACT

The brain's ability to alter its functional and structural architecture in response to experience and learning has been extensively studied. Mental stimulation might serve as a reserve mechanism in brain aging, but macrostructural brain changes in response to cognitive training have been demonstrated in young participants only. We examined the short-term effects of an intensive memory training program on cognition and brain structure in middle-aged and elderly healthy volunteers. The memory trainers completed an 8-week training regimen aimed at improving verbal source memory utilizing the Method of Loci (MoL), while control participants did not receive any intervention. Both the memory trainers and the controls underwent magnetic resonance imaging (MRI) scans and memory testing pre and post 8 weeks of training or no training, respectively. Cortical thickness was automatically measured across the cortical mantle, and data processing and statistical analyses were optimized for reliable detection of longitudinal changes. The results showed that memory training improved source memory performance. Memory trainers also showed regional increases in cortical thickness compared with controls. Furthermore, thickness change in the right fusiform and lateral orbitofrontal cortex correlated positively with improvement in source memory performance, suggesting a possible functional significance of the structural changes. These findings demonstrate that systematic mental exercise may induce short-term structural changes in the aging human brain, indicating structural brain plasticity in elderly. The present study included short-term assessments, and follow-up studies are needed in order to assess whether such training indeed alters the long-term structural trajectories.

© 2010 Elsevier Inc. All rights reserved.

Introduction

Plasticity and reorganization of neural systems have recently been studied in both animals and humans (Buonomano and Merzenich, 1998; DeFelipe, 2006; Jones et al., 2006; Martin et al., 2000). Advances in neuroimaging techniques have enabled tracking of behavioral changes to alterations in specific brain regions *in vivo* (Draganski and May, 2008; Haier et al., 2009). Accumulating evidence suggests that the brain's potential to adapt and change is life-long (Johansson, 2004; Pascual-Leone et al., 2005). However, the mechanisms by which such alterations are effectuated are poorly understood, and little is known about their functional specificity. Structural grey matter (GM) changes measured by MRI have been reported in young subjects from days to months after learning to juggle (Draganski et al., 2004; Driemeyer et al., 2008). This finding was also replicated in older subjects, suggesting intact neuroplasticity in advanced age (Boyke et al., 2008). The same

group also reported structural alterations in students in response to extensive studying (Draganski et al., 2006). Thus, training visuo-motor skills and learning abstract information seems to induce structural brain changes in both young and older subjects.

To answer questions of domain specificity, it is important to address different functional domains. Memory, a spectrum of cognitive functions inextricably coupled to everyday function, is associated with increasing concern among the aging population (Anderson and McConnell, 2007; National Council on the Aging, 2000), and memory complaints are reported by up to 50% of adults aged 64 and over (Reid and MacLulich, 2006). Furthermore, aging is associated with both progressive structural alterations (Fjell et al., 2009b; Jernigan et al., 2001; Raz et al., 2004, 2007; Salat et al., 2004; Walhovd et al., 2005, 2009; Westlye et al., 2009, *in press*-a,b) and functional decline across multiple cognitive domains (Mahncke et al., 2006; Raz and Rodrigue, 2006). Population studies have indicated that mental exercise may slow the rate of cognitive decline (Valenzuela and Sachdev, 2006a) and decrease the risk of dementia (Valenzuela and Sachdev, 2006b). Hence, the role of mental exercise has gained much attention in preventive aging research (Rebok et al., 2007). However, knowledge

* Corresponding authors. Center for the Study of Human Cognition, Department of Psychology, University of Oslo, Oslo 0317, Norway.

E-mail address: andreas.engvig@gmail.com (A. Engvig).

about protective effects of mental exercise is sparse, and it is not known whether engaging in cognitively stimulating activities changes aging-related structural cerebral trajectories.

Most behavioral studies of memory improvement have been strategy-based, focusing on learning a specific mnemonic technique known as the MoL (Bower, 1970). Studies employing this technique have found associated changes in brain activation and neurochemistry (Kondo et al., 2005; Nyberg et al., 2003; Valenzuela et al., 2003). However, no experimental study so far has been conducted to investigate whether such memory training can induce macro-structural changes in the brain. Further, visuo-motor training has been shown to induce macro-structural changes in elderly, but for cognitive training this has been demonstrated in young only. The aim of the present MRI study was to determine whether engaging in an 8-week memory training program would improve memory performance and induce regional changes in longitudinal cortical thickness trajectories in middle aged and older adults.

Methods

Sample

Table 1 summarizes baseline demographic and neuropsychological characteristics of the participants included in the analyses. Fig. 1 gives an overview of the recruitment and group assignment process. Volunteers were recruited through a local newspaper ad and screened by a structured interview before inclusion. All included participants reported to be right-handed, native Norwegian speakers, not concerned about their own memory function, not using medications known to interfere with cognitive function (including benzodiazepines, antidepressants or other central nervous agents) and having no diseases known to affect the central nervous system (CNS), including thyroid disease, multiple sclerosis, Parkinson's disease, stroke, severe hypertension or diabetes.

48 volunteers were recruited and assessed for eligibility. As outlined in Fig. 1, 45 participants were randomly assigned to either of two groups: (1) an intervention group participating in an 8-week intensive memory training program, or (2) a control group serving as passive controls. The groups were carefully matched by sex, age and education level. 23 trainers and 20 of the 22 controls completed the study and were paid for participation. To minimize the influence of subclinical degenerative conditions, the following exclusion criteria were employed: Mini Mental Status Exam (MMSE) <26 (Folstein et al., 1975), Geriatric Depression Scale (GDS) >11 (Yesavage et al., 1983)

Table 1

Characteristics of completing participants at baseline ($n = 42$).

	Memory training (12F/10 M)			Control (11F/9 M)			p-value*
	M	SD	Range	M	SD	Range	
Age	61.3	9.4	42–76	60.3	9.1	42–77	0.74
Education	15.1	1.9	12–18	15.6	1.8	12–18	0.48
IQ	118.0	8.9	99–132	118.8	9.2	92–130	0.80
MMSE	29.0	1.0	26–30	29.1	0.9	27–30	0.63
GDS	1.6	1.9	0–6	1.5	2.2	0–8	0.84
TMT A	36.8	15.8	23–98	33.0	13.8	16–63	0.42
TMT B	80.6	34.8	48–209	70.7	32.7	37–159	0.35
Digit-symbol	65.8	15.5	36–96	67.7	14.1	41–89	0.68
Rey-O 30 min recall	19.1	6.7	7–31	21.1	5.7	12–32	0.30
CVLT 1–5 total	50.0	9.1	31–63	52.3	10.7	33–70	0.57
CVLT long delay	11.6	2.3	7–15	11.9	3.6	5–16	0.89

Note. IQ from derived from WASI matrices and vocabulary sub tests. MMSE, Mini Mental State Exam. GDS, Geriatric Depression Scale. TMT A and B, trail making test part A and B, seconds; Digit-symbol from WAIS-R; Rey-O, 30 minutes delayed recall of the Rey-Osterreith complex figure test; CVLT 1–5 total, score with intrusions subtracted; CVLT long delay, score on 20 minutes delay recall with intrusions subtracted.

* p-values are two-tailed and based on independent samples *t*-tests with training or control as grouping variable.

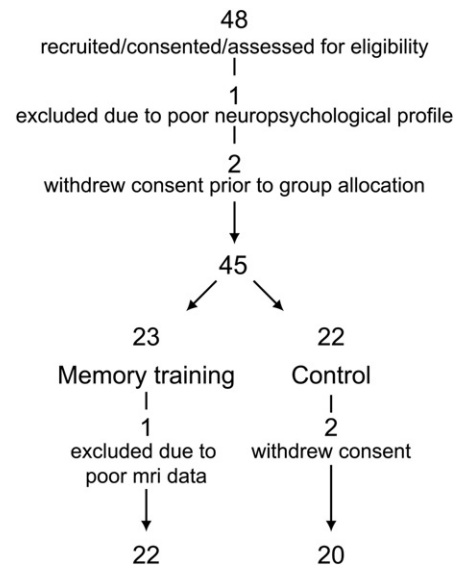


Fig. 1. Flow chart of volunteer recruitment.

and IQ <85, derived from the vocabulary and matrices sub tests in the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999). One participant only filled out the GDS form after completion of the program, however the score (5.33) was below the exclusion criterion. Further, participants with a total learning and/or recall score on the California Verbal Learning Test (CVLT-II) (Delis et al., 2000) of 2 standard deviations (SD) or more below population norm (Delis et al., 2000) were also excluded. One participant was excluded on the basis of the CVLT score. Furthermore, all MRI scans were subjected to a radiological evaluation by a neuroradiologist, and were required to be deemed free of significant injuries or conditions (e.g. signs of brain tumors or stroke). None of the participants were excluded on the basis of this. One subject was excluded due to inadequate MRI data quality (see MRI image processing/reconstruction).

Cognitive assessments and the memory training program were conducted at the Center for the Study of Human Cognition at the Department of Psychology, University of Oslo. All participants gave informed consent, and the study was approved by the Regional Ethical Committee of South Norway (REK-Sør). At the start of the project-time point one (tp1)—all participants were tested on a series of neuropsychological and behavioral tests (see above and below), and underwent an MRI scan. The following week, the memory training group started an 8-week training scheme, while the controls were instructed to continue living as usual (participants were informed of their group assignment after the neuropsychological and behavioral testing). At time point two (tp2) approximately 9 weeks later and after completion of the memory training, all participants were re-examined on selected behavioral tests and underwent a second MRI scan. Mean interval between the two scanning sessions was 65.3 days (SD = 8.0, min: 54 days, max: 89 days).

Memory training program

An 8-week training program was designed to improve serial verbal recollection memory by implementing the visualization mnemonic technique MoL (Bower, 1970). MoL involves learning 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. After acquisition of a loci route, the to-be-remembered information is linked to the various loci at the time of encoding. At test, the landmarks are mentally revisited in serial order, and the information associated with each locus is retrieved. This method is shown to substantially improve serial recall in older adults (Kliegl et al., 1990). In the current intervention, each training week consisted of a 1-h class room session and 4 days of

home work exercises. The class room sessions consisted of small groups lead by an instructor. 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. The first two sessions focused specifically on demonstrating and troubleshooting the use of the MoL, respectively.

Eight standardized exercises were made to observe the use and improvement of the memory training in-class. The exercises were administered at the end of each session throughout the course. Each exercise consisted of three lists of 10, then 20 and, finally, 30 concrete words which the participants were told to memorize in sequential order using the MoL. Participants were given a maximum time of five, seven and ten minutes to complete the respective lists. Participants had to successfully recall a 10-words list in order to proceed with a 20-word list, and so on. Serial recall of the 30-word list was considered difficult without applying the MoL successfully. The 32 home exercises consisted of lists of concrete words to be remembered. For motivational purposes, the home exercises also consisted of lists with names of different countries, words in foreign languages, names of American presidents, Roman emperors, flower names and the names of Norwegian ministers. However, each exercise 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. Two of the home assignments also required the participants to remember names and faces by using the MoL. Home work effort was balanced across weeks.

Practice and test stimuli were drawn from concrete words in 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>). Word length (mean 5.51, range 3–9, SD = 1.39) and frequency (mean 49.5 out of 18.5 million OC occurrences, range: 1–198, SD = 43.1) were balanced across lists. English translations of one in-class exercise and one home exercise are provided as [Supplementary Material](#). The complete material from the memory training program can be obtained in Norwegian upon request. We calculated the participation rate by registering the number of the maximum 40 (eight in-class and 32 at home) exercises each volunteer performed, allowing no less than 50% participation rate (mean 96.9%, SD = 9.4 Range [57.5–100.0]). One participant had a low participation rate (57.5%). We decided to include this participant, who showed adequate improvement in both in-class exercises and source memory re-testing. The other participants had a rate above 82% (range: 82.5–100.0).

Memory assessment

Word recognition and source memory function was assessed at both time-points using two versions of a computerized test developed for the current study. Briefly, the test consisted of one practice block and four experimental blocks. Prior to each block participants were instructed to try to remember the words that would be presented on the screen, as well as the sequential order in which they were presented. In each experimental block a series of 15 consecutive words were presented on screen, each for 1 s, with 4-s intervals of blank screen between each word. Immediately after the presentation of the word list, participants were informed on screen that they would again be presented a series of words (a semi-randomized mix of 15 previously presented and 15 new words). They were instructed to indicate by button presses within 5 s after each word presentation whether the word was part of the list they had just seen (recognition task). If the word was rejected as new, the next word was presented. If the participants indicated that they had previously seen the word, they were prompted to indicate at self pace whether it was among the first 5, the middle 5, or the last 5 words on the list (source memory task). The practice blocks were identical to the experimental blocks,

with the exception that only 6 words were presented (these data were not used in the analysis). With four experimental blocks of 15 words each, maximum recognition and source-memory score for each time-point was 60. Word lists used in the pre- and post-test were matched with regard to word frequency and number of letters.

MoL training is shown to improve one's ability to recall the order of concrete words (source memory) (Verhaeghen et al., 1992), while recognition memory is not considered to be improved by recollection training (Jennings et al., 2005). Thus, we used the difference in proportion of correct source memory hits (correct source memory judgements/recognition hits, for each tp) between tp1 and tp2 as a measure of memory improvement specifically targeted by MoL. A ratio of 1 at either tp implies that a participant correctly recalled the order of all correctly recognized words. One participant in the control group did not perform the post-test and was thus excluded from the behavioral analyses.

MRI acquisition

MRI acquisitions were performed at Oslo University Hospital, Rikshospitalet using a 12 channel head coil on a 1.5-Tesla Siemens Avanto scanner (Siemens Medical Solutions, Erlangen, Germany). The pulse sequence used for morphometric analyses were two repeated 3D T1-weighted Magnetization Prepared Rapid Gradient Echo (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. Each scan took 7 min 42 s. Each volume consisted of 160 sagittal slices with an in plane resolution of 1.25 × 1.25 mm and 1.20 mm slice thickness. The acquisitions were automatically aligned to 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). After acquisition the data were deidentified and exported to Linux workstations for processing and analyses.

MRI image processing/reconstruction

Data processing and analyses were performed at the Neuroimaging Analysis Lab at the Center for the Study of Human Cognition, University of Oslo, with use of additional computing resources from the Titan High Performance Computing facilities at the University of Oslo (<http://hpc.uio.no/index.php/Titan>). All scans were manually 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 (SNR) and thickness measurement reliability. FreeSurfer (<http://surfer.nmr.mgh.harvard.edu>) was used for reconstruction of the brain's surface to compute cortical thickness at each vertex using a semi-automated approach described in detail elsewhere (Dale et al., 1999; Dale and Sereno, 1993; Fischl and Dale, 2000; Fischl et al., 1999a,b; Salat et al., 2004). The thickness measurements were obtained by reconstructing representations of the gray/white matter boundary and the pial surface (Dale et al., 1999; Dale and Sereno, 1993) and then calculating the distance between those surfaces at each vertex across the cortical mantle. This method uses both intensity and continuity information from the entire three-dimensional MR volume in segmentation and deformation procedures to produce surface mapped representations of cortical thickness. The surface is created using spatial intensity gradients across tissue classes and is therefore not simply reliant on absolute signal intensity. The surfaces produced are not restricted to the voxel resolution of the original data and the procedure is thus capable of detecting submillimeter differences between groups (Fischl and Dale, 2000). The measurement technique has been validated via histological (Rosas et al., 2002) and manual measurements (Kuperberg et al., 2003). All reconstructed data were visually checked for segmentation accuracy at each time point. One participant in the training group was excluded from analyses due to MRI subject

motion artifacts, bringing the total sample down to 42 subjects. No manual interventions with the MRI data were performed.

A longitudinal processing scheme implemented in FreeSurfer was used in the present study. This aims to incorporate the subject-wise correlation of longitudinal data into the processing stream to reduce the measurement noise and ensure non-biased analysis of changes in cortical thickness. In short, the data from both time points are initially processed cross-sectionally. Next, a base template is made from an average of the two time points. The first and second time point is registered to this base template to ensure non-biased analysis with regards to any of the time points. This scheme significantly increases thickness measurement reliability compared with the default scheme (Han et al., 2006). Each individual's thickness map from the longitudinally processed tp1 was then subtracted from the thickness map from the longitudinally processed tp2 in a vertex-wise manner. The resulting individual difference maps were resampled, mapped to a common surface, smoothed with a Gaussian kernel with a full width of half maximum (FWHM) of 30 mm and submitted to statistical analyses. The relatively wide smoothing kernel was chosen in order to reduce noise in the thickness measurements and increase sensitivity and validity of statistical analysis. Morphometry-cognition relationships are often not very focal, at least not in a cognitive interventional study, and we chose a kernel size accordingly. Note, however, that this also reduced the spatial resolution of the thickness maps (Han et al., 2006).

Statistical analyses

Independent samples *t*-tests were performed to compare baseline demographical and neuropsychological data between the memory training and control group. Repeated measures analysis of variance (ANOVA) and independent-samples *t*-tests were performed to compare memory improvement between groups. The memory training and control group were compared to reveal any baseline differences in cortical thickness by fitting a general linear model (GLM) of the effect of group on thickness at tp1 at each vertex across the surface.

We hypothesized that the present memory training scheme would be associated with regional structural changes. However, to test for global effects and to compare our data with previous literature, we first estimated global effects as indexed by changes in total brain volume (TBV) and total cortical GM volume (TCV). We also calculated the mean thickness of all measured vertices across the cortical mantle, mean cortical thickness (MCT). TCV and MCT were calculated from the cortical surface reconstructions described above. TBV was estimated by a validated (Jovicich et al., 2009) automated procedure for volumetric measures of different brain structures provided in FreeSurfer. The procedure automatically assigns a neuroanatomical label to each voxel in an MRI volume based on probabilistic information automatically estimated from a manually labeled training set (Fischl et al., 2002). In the present study TBV was defined as the sum of the volume of brain structures identified by FreeSurfer, after correcting for partial-volume effects. This includes cerebral and cerebellar GM and WM and subcortical structures, but not the ventricles, CSF and dura. To rule out differences in intracranial volume (ICV) as a potential confound, we estimated ICV at tp1 and tested for group differences using an independent samples *t*-test.

Group differences in baseline TBV or TCV were tested using independent samples *t*-tests. Effects of memory training on TBV and TCV were tested by repeated measures ANOVA and independent-samples *t*-tests on mean TBV atrophy rate and TCV change between groups.

To compare the estimated longitudinal brain changes in our data with previous literature, we calculated annualized TBV atrophy rates by scaling the TBV rate for each group. Since we did not expect the memory training scheme to induce large global structural changes, the global analysis was mainly performed in order to validate the

estimated rates of change in our data by comparing the current dataset to previous published studies.

Secondly, vertex-wise GLM was performed to test for training-related regional changes in cortical thickness. We investigated group \times time interactions on thickness, testing whether the mean paired differences in the intervention group differed from that in the control group, as such modeling specific effects related to the memory training. The vertex-wise analyses were performed using group as class variable and vertex-wise difference maps (tp2-tp1) as data matrix without any covariates, which reduces to vertex-wise *t*-tests.

For the group \times time interaction analysis, we controlled the family-wise error (FWE) rate by the means of Z Monte Carlo simulations (Forman et al., 1995) as implemented in FreeSurfer (Hagler et al., 2006; Hayasaka and Nichols, 2003). The method is based on AFNI's AlphaSim (Ward, 2000). The data was tested against an empirical null distribution of maximum cluster size across 10,000 iterations synthesized with an initial cluster forming threshold of $p < 0.001$, thus yielding clusters fully corrected for multiple comparisons across the surface. Since cluster size inference evaluates clusters based on cluster extent, it may penalize spatially localized yet reliable effect sites, e.g. within small structures as the entorhinal cortex. Thus, we additionally assessed the reliability of regional clusters in the uncorrected data using a split-half strategy. The 22 participants in the training group were sorted according to sex, age and MoL performance (mean percentage of sequentially recalled words at the seventh and eighth in-class session). Then the first on the list was allotted to split-half group one ($n = 11$), number 2 on the list to split-half group two ($n = 11$) and so on. The mean of the individual tp2-tp1 difference maps from the two subgroups were then compared independently with the control group. Furthermore, we split the control group and performed four independent GLMs using all four split groups to further ensure independence of the data sets. The results from these cross-validated split-half analyses are presented in Supplementary Fig. 1. Thus, in addition to the correction for multiple comparisons using cluster size-inference, we identified regionally consistent effects based on split-half analysis. Consistent areas between split-halves were found quantitatively by mapping each split-half map to a common template brain. Next, we used custom Matlab procedures to identify vertices that were significant in all uncorrected split analyses and mapped these to the template brain for visualization and estimation of spatial extent (area). Effect sites identified using both methods as well as the uncorrected spatial *p*-distribution are reported. Histograms of the *p*-value distributions across the surfaces are plotted for both hemispheres to show the global effects.

For post-hoc analysis of brain/behavior interactions, we thresholded the statistical maps from the main interaction analysis at $p < 0.01$, and calculated mean thickness at each time-point within the consistent regions. These were submitted to linear regressions with behavioral measures as dependent variables. Specifically, we regressed thickness at both time points and thickness change in regions showing training-related thickness changes on source memory performance and time spent on home exercises to test for possible relations with performance and practice rate. Furthermore, we regressed thickness change in the regions identified in the main analysis with age. The post hoc analyses were applied independently to the group \times time analysis which modeled the effects of group and time on thickness without any covariates (e.g. memory performance). Thus the present brain-behavior analyses did not include several dependent tests under the same null-hypothesis (circular inference or 'double dipping') (Kriegeskorte et al., 2009; Vul et al., 2009), but rather independent tests of two separate hypotheses: (1) memory training induces cortical thickness changes and (2) the rate of cortical thickness change correlates with rate of training-related improvement or time spent on exercise.

As we hypothesized that source memory improvement per se was related to thickness change, we also correlated thickness change with memory improvement while adjusting for baseline memory score.

This accounts for potential influence of baseline memory score on the relationship between memory improvement and cortical thickness change. The rationale for regressing out memory score at tp1 was that we wanted to study the effect of change per se, regardless of the initial memory performance. If the raw scores for memory improvement (tp2 score–tp1 score) were used, a change in recall from 12 to 14 would be equal to a change from 4 to 6. Regressing out tp1 memory score was therefore done to correct for this.

To assess the validity of the observed cortical thickness changes in the control group we included nine additional healthy middle-aged males whom were currently available from an ongoing longitudinal study on healthy aging at our lab. The participants were scanned with the same MRI protocol ($n=9$, mean age = 53.4 years) at two time points. The participants of this replication sample were screened with the same questionnaire, and reported not be concerned about their own memory function. We compared monthly cortical change in each of the reported regions between the control and replication groups.

Results

Memory performance changes

Table 1 summarizes neuropsychological characteristics and results from independent samples t -tests comparing the two groups at tp1. No significant group differences were found at tp1. The mean source memory task scores for the training and control group were 0.52 (Standard Error (SE) = 0.03) and 0.57 (SE = 0.02), respectively, at tp1, and 0.73 (SE = 0.03) and 0.62 (SE = 0.03), respectively, at tp2. Repeated measures ANOVA revealed a significant group \times time interaction (Greenhouse–Geisser corrected, $F(1, 39) = 14.8, p < 0.001$).

Group means and standard deviations for the recognition and source memory scores at each time point are provided in Table 2. Briefly, the results revealed a specific increase in number of correct source memory judgements in the training group ($p < 0.001$). While no changes were found for number of correct recognitions, the training group also exhibited a decreased number of false alarms after the training regimen ($p = 0.006$).

Longitudinal analyses confirmed that participants in the training program (mean difference = 0.21, SD = 0.14) showed a significant improvement in the source memory task as measured by the source/recognition ratio (paired t -test, $t = 7.04, df = 21, p < 0.0001$), while for the control group a trend improvement on the same scores (mean difference = 0.05, SD = 0.12) failed to reach statistical significance (paired t -test, $t = 1.73, df = 18, p = 0.10$). An independent samples t -test on the difference in source memory improvement showed a significant difference in improvement between groups ($t = 3.85, p < 0.001$).

Table 2
Group means and standard deviations for memory performance at tp1 and tp2.

	Baseline (tp1)		Retest (tp2)		p -value*
	M	SD	M	SD	
<i>Memory training (12F/9M)</i>					
Recognition, correct hits	54.0	3.99	53.6	3.81	0.736
Recognition, false alarms	3.13	5.23	1.27	3.07	0.006
Correct source memory judgements	28.1	9.04	39.1	9.82	<0.001
Source/recognition, ratio	0.512	0.146	0.727	0.160	<0.001
<i>Controls (10F/9M)</i>					
Recognition, correct hits	52.3	5.53	54.2	3.49	0.097
Recognition, false alarms	2.21	2.49	2.00	1.89	0.736
Correct source memory judgements	29.8	6.88	33.5	6.62	0.063
Source/recognition, ratio	0.568	0.106	0.618	0.115	0.101

Note. 'Recognition, false alarms' denotes new words that are categorized as previously displayed.

* p -values are based on paired samples t -tests.

Global brain changes

The mean global measures at tp1 and tp2 are reported for both groups in Supplementary Table 1. Independent samples t -tests revealed no differences in TBV ($t = 1.10, p = 0.28$) or ICV ($t = 0.46, p = 0.65$) between groups at tp1. Repeated measures ANOVA on TBV and group, testing whether mean change in TBV differed between groups, revealed a trend group \times time interaction, but failed to reach statistical significance (Greenhouse–Geisser corrected, $F(1, 40) = 2.88, p = 0.098$). Annualized rates of whole-brain atrophy (TBV) were estimated for the training (mean = 0.30%, SD = 3.25) and control group (mean = -1.51%, SD = 3.24), respectively. Positive values indicate a relative growth and negative values a relative atrophy. One-sample t -tests indicated a trend relation towards whole-brain atrophy ($t = -2.1, p = 0.051$) in the control group. The training group's atrophy rate did not differ significantly from 0 ($t = 0.43, p = 0.67$). Independent samples t -test on the annualized TBV rates revealed a trend difference between groups ($t = 1.81, p = 0.078$), but the results failed to reach statistical significance.

Independent samples t -tests revealed no significant baseline differences in TCv ($t = 0.78, p = 0.44$) and MCT ($t = 0.035, p = 0.97$). Independent samples t -tests on change in TCv ($t = 0.79, p = 0.44$) and MCT ($t = 0.93, p = 0.36$) showed no significant differences between groups.

Regional brain changes

We found no significant regional differences in cortical thickness between groups at tp1. Fig. 2 shows the results from the group \times time interaction analysis. The results revealed cortical thickening in the training group compared with controls in the right insula (R1) ($p < 0.05$, corrected for multiple comparisons, cluster size = 298 mm², Talairach coordinates at peak vertex [$X = 36.3, Y = -0.6, Z = 6.1$]). Fig. 3 shows the results from the split-half analysis, which further suggested significant ($p < 0.01$, uncorrected) training-related cortical thickening in the left lateral orbitofrontal cortex (L1), and the right lateral orbitofrontal cortex (R2) and fusiform cortex (R3). Further splitting of the control group revealed equivalent patterns across all analyses, but reduced significance (Supplementary Fig. 1). No areas of significant thinning in the training group compared with controls were found. The split half validated spatial p -distribution maps from the GLM testing group \times time interactions are displayed in Fig. 4A and C.

As evident from Figs. 2 and 4B, the training group showed a significant cortical thickening in the right insular and fusiform cortices. Also, the control group exhibited a significant thinning in the left lateral orbitofrontal cortex and the insular and the fusiform cortex in the right hemisphere. A negative correlation between age and right insular thickness change were found in the training group (Pearson's $r = -0.45, p = 0.037$), indicating decreasing training-related cortical thickening with increasing age. No other correlations between age and changes in cortical thickness were found.

Fig. 5 shows the global distribution of p -values from the interaction analysis for both hemispheres. The distribution is shifted to the right (relative thickening in the intervention group compared with controls). Two-tailed one-sample t -tests confirmed distribution means > 0 in the left [mean = 0.22, $t = 230, p < 0.0001$] and right [mean = 0.29, $t = 197, p < 0.0001$] hemisphere, respectively.

Brain-behavior relationships

We found a significant positive correlation between memory improvement and cortical thickening in right fusiform cortex (R3) (Pearson's $r = 0.37, p = 0.019$). Cortical thickening in the right lateral orbitofrontal cortex (R2) and memory improvement showed a trend ($r = 0.25, p = 0.12$) relation. Statistically adjusting for score at tp1

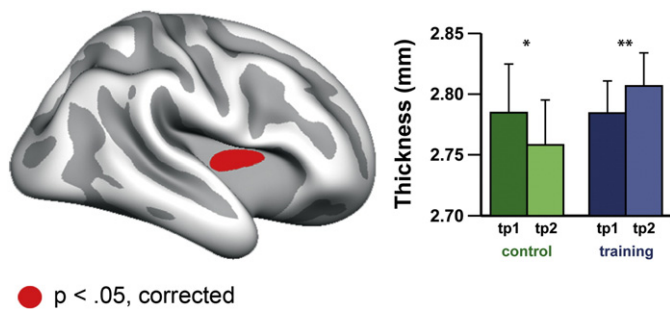


Fig. 2. Memory training increases cortical thickness in right insula (R1). Changes in right insular cortical thickness in response to memory training estimated by comparing the training and control groups by means of general linear models (group \times time interaction). The results are corrected for multiple comparisons, using cluster size-inference. The corrected p -value cluster ($p < 0.05$, cluster size = 298 mm²) is mapped to an inflated template brain for better visualization of the insula region. To the right, bar plots for mean R1 cortical thickness at scan 1 and 2 are displayed. Error bars represent 1 Standard Error (SE). Asterisks (*) and (**) indicate $p < 0.05$ and $p < 0.01$ based on two-tailed paired samples t -tests for cortical thickness before and after training, respectively.

strengthened the correlation between memory improvement and cortical thickening in R3 ($r = 0.39$, $p = 0.012$) and R2 ($r = 0.32$, $p = 0.041$), respectively. Fig. 4D illustrates the relationship between thickness change in these areas and change in memory performance controlled for baseline memory score for each participant. Additionally, in the left hemisphere, lateral orbitofrontal (L1) thickness at tp1 ($r = 0.35$, $p = 0.023$) and tp2 ($r = 0.41$, $p = 0.008$) correlated positively with the source/recognition memory ratio at tp2. This indicates that a thicker cortex in the L1 area both at baseline and follow-up correlated positively with memory performance at follow-up. There were no group differences in L1 thickness at tp1 (two-tailed independent samples t -test, $p > 0.1$). No other significant correlations

between thickness and memory performance, time spent on each exercise or participation rate were found.

Replication sample

Monthly cortical change for the control and replication sample was calculated for L1 [mean = -0.017 (SD = 0.028)% and -0.014 (0.038)%, respectively], R1 [-0.015 (0.024)% and -0.012 (0.045)%, respectively], R2 [-0.012 (0.069)% and -0.006 (0.022)%, respectively], and R3 [-0.024 (0.375)% and 0.009 (0.033)%, respectively]. Independent samples t -tests revealed no significant differences in regional cortical thinning between the control group and the replication sample for L1, R1 and R2 ($p > 0.7$). The control group exhibited a more pronounced thinning in the R3 (right fusiform) area compared with the replication sample ($p < 0.05$).

Supplementary Fig. 2 shows within-group changes (tp2–tp1) in cortical thickness ($p < 0.05$, uncorrected). Briefly, nine clusters of regional thinning and a minor area of thickening in the left superior frontal gyrus were found in the control group. The training group exhibited six areas of thickening and one area of thinning in the left middle temporal gyrus. None of these effects survived corrections for multiple comparisons and should thus be interpreted with caution.

Discussion

There were three main findings in the present study: (1) Intensive source memory training specifically improved memory performance on a source memory task, (2) memory training induced longitudinal short-term regional effects on cortical thickness, indicating that the intervention influenced the age-related trajectories, and (3) regional cortical thickening was positively correlated with memory improvement. Thus, degree of cortical change was directly related to improvement rate in source memory performance. This is the first study

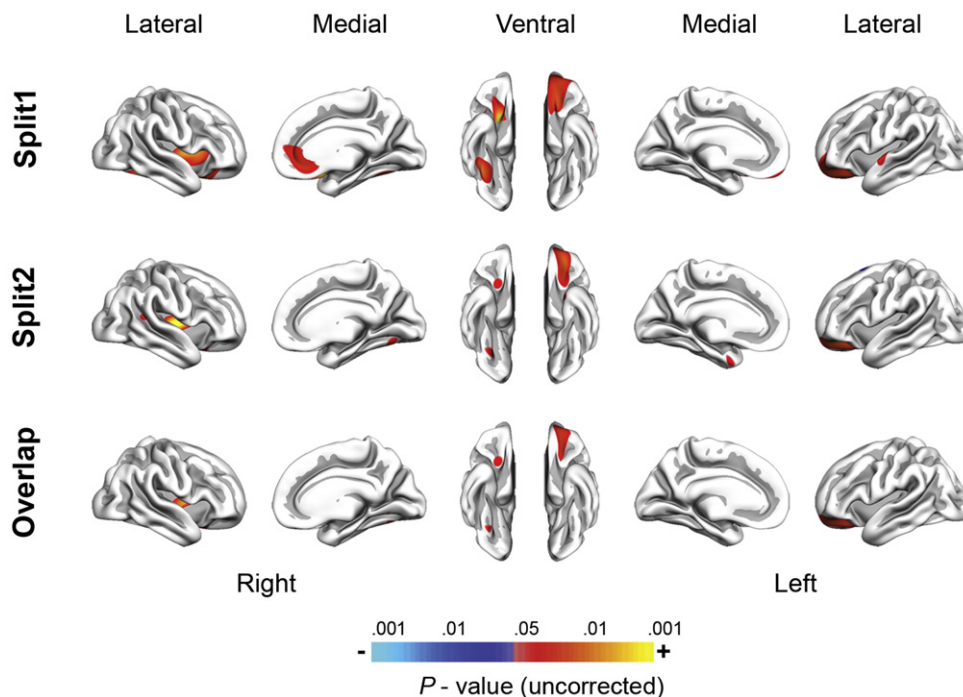


Fig. 3. Split half validation. The surface-mapped statistical p -maps show relative changes in cortical thickness between the two memory training subgroups and the control group (split 1 and split 2) estimated by a general linear model analysis using the split-half approach. Blue color indicates thinning in the training subgroups relative to the control, whereas red indicate thickening relative to the control group. The p -value maps are thresholded at $p < 0.05$ (uncorrected). Row three displays consistent overlaps of significant vertices between the two split-halves: right hemisphere: insula (1373 overlapping vertices, area = 525.2 mm²), lateral orbitofrontal cortex (349 overlapping vertices, area = 168.3 mm²) and fusiform gyrus (149 overlapping vertices, area = 90.8 mm²); left hemisphere: lateral orbitofrontal cortex (1293 overlapping vertices, area = 884.5 mm²). The clusters in row three are represented by the lowest p -value (least significant) among the two above maps. All overlapping vertices indicated a relative thickening in the training group. No overlapping vertices with the inverse relation (relative thickening in the control groups) were found.

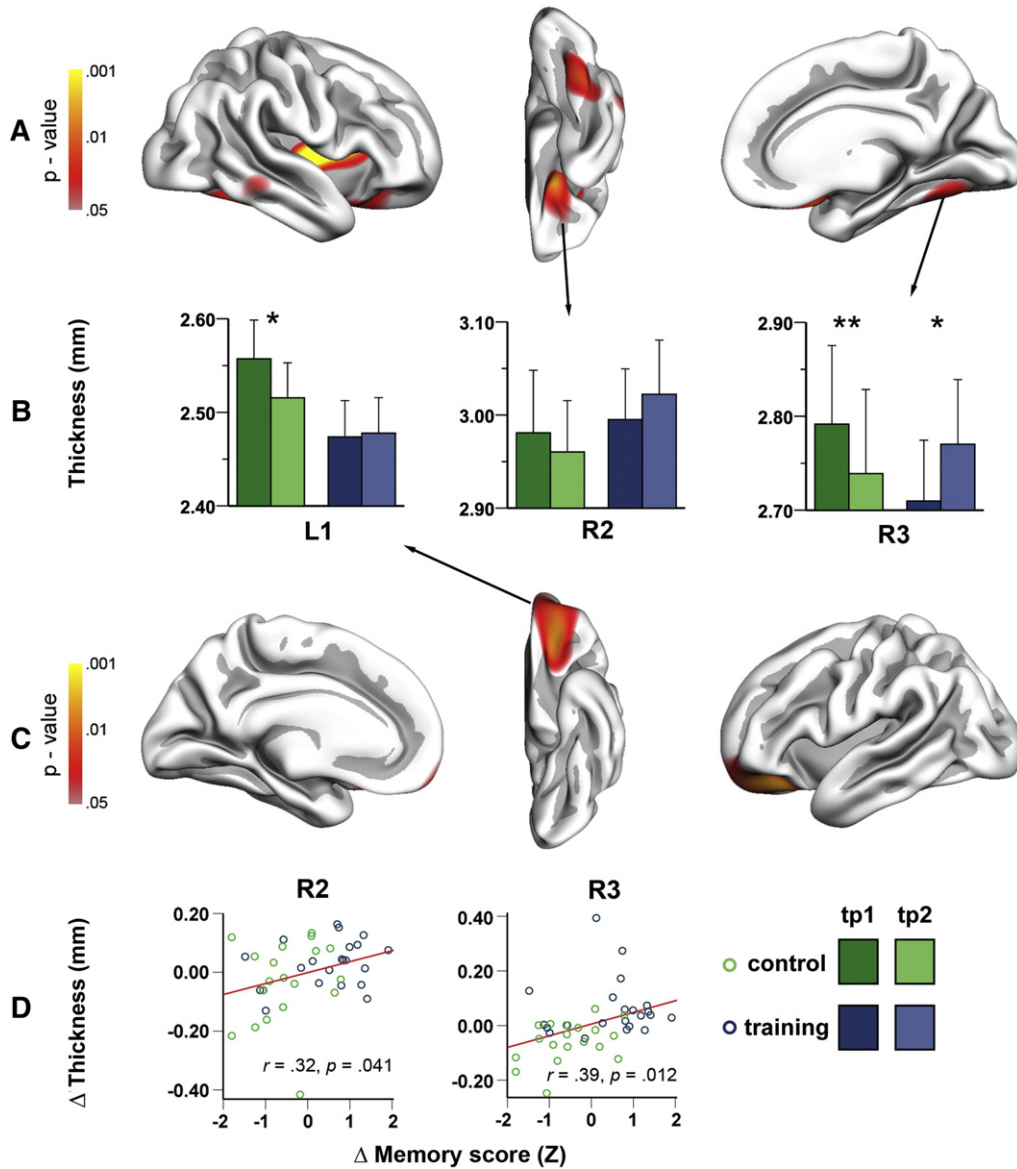


Fig. 4. Memory training regionally increases cortical thickness. Changes in cortical thickness in response to memory training estimated by comparing the training and control groups by means of general linear models (group \times time interaction). Only areas replicated in the split half validation are reported. The statistical p -maps are color coded and mapped to a semi-inflated common template brain for visualization of effects buried in sulci. The warm color gradient denote relative cortical thickening in the training group. No areas of relative thinning with respect to the training group were found. The displayed maps are thresholded at $p < 0.05$ for visualization, and reported data for bar- and scatter plots at $p < 0.01$. (A) Right hemisphere, lateral, ventral and medial view. (B) Bar plots for mean cortical thickness at scan 1 and 2 from three consistent areas found in the split-half analysis: Left hemisphere: L1 = lateral orbitofrontal; Right hemisphere: R2 = lateral orbitofrontal, R3 = fusiform. Right insula (R1) is reported in Fig. 2. Error bars represent 1 Standard Error (SE). Asterisks (*) and (**) indicate $p < 0.05$ and $p < 0.01$ based on two-tailed paired samples t -tests for cortical thickness before and after training, respectively. (C) Left hemisphere, medial, ventral and lateral view. (D) Scatter plots showing cortical thickness change as a function of change in memory performance, independent of initial task performance for the R2 and R3 areas.

to demonstrate altered short-term trajectories on cortical thickness in elderly in response to an intensive memory training intervention. The main findings will be discussed in detail below.

The specific improvement in source memory scores after memory training was in accordance with previous studies aimed at improving memory performance (Verhaeghen et al., 1992). This indicates that the training program successfully improved a task-specific cognitive domain, at least temporarily. Follow-up studies are needed in order to assess the long-term effects of the memory training regimen, both with respect to how the increased memory performance transfers to daily-life functioning and ultimately to which degree such training may serve as a reserve mechanism and protect against cognitive deterioration. Although previous studies have found that memory training improves memory performance, this has not been studied in relation to brain structure.

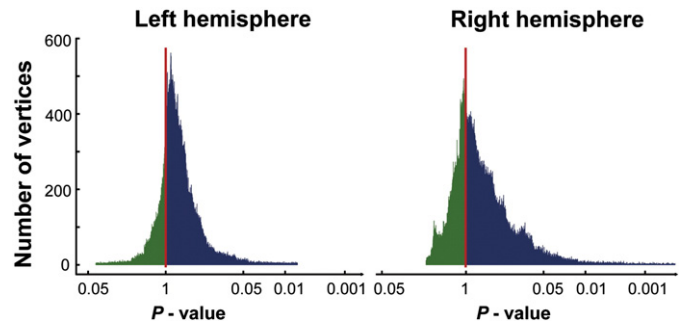


Fig. 5. Global distributions of p -values. The figure shows the number of vertices for any given p -value from the cortical differences maps illustrated in Fig. 4 across the surface. The left side of each chart (green color) represents a relative cortical thickening in the control group. The right side (dark blue color) represents a relative thickening in the training group. The distribution of p -values is shifted to the right, mostly so in the right hemisphere.

Positive effects of interventions may manifest as cortical thickening or reduced age-related cortical thinning. Thus, longitudinal effects of the memory training regimen on cortical structure may be detected by testing the group by time interactions on cortical thickness. We found significant group \times time interactions on cortical thickness in the right insular cortex, indicating cortical thickening or reduced thinning in the memory training group compared to the controls. Follow-up analyses revealed that while the training group showed significant thickening in this area, the control group showed significant thinning.

Converging evidence supports insular involvement in interoception, emotional integration and awareness (Craig, 2009). Although the specific involvement of the insular cortices in memory formation remains unclear, the right insula is reported to play a role in controlling goal-directed behavior (Dosenbach et al., 2007). Sridharan and colleagues (2008) propose that the right anterior insula is part of a distinct network playing a critical role for switching between distinct brain networks, engaging the brain's attentional, working memory and higher-order control processes while disengaging other systems that are not task-relevant. It is possible that such a suggested network played a role in this study, requiring the participants to encode, recall and finally examine the recall performance, almost daily for 8 weeks. However, the fact that change in right insular thickness did not correlate with change in memory performance might indicate that thickening in this region reflect processes not associated with source memory improvement.

The brain-behavior analyses revealed that improvement of memory performance after 8 weeks of intensive memory training was associated with cortical thickening in the right fusiform cortex and right lateral orbitofrontal cortex, independent of initial task performance. Thus, participants showing the largest improvement in memory performance also showed the strongest thickening in these areas. Further, cortical thickness in the left lateral orbitofrontal cortex was associated with re-test scores on a source memory task. Specifically, thicker cortices both before and after training was associated with a higher re-test scores. Although effects were observed bilaterally, they were more pronounced in the right hemisphere. Although the reasons for this apparent lateralization is unclear, the right hemisphere is essential for sustaining attention (Posner and Petersen, 1990) and visuospatial processing (Corballis, 2003), domains targeted by the current memory training program.

Both the fusiform and lateral orbitofrontal cortices are implied in memory function. The fusiform gyri are essential for object perception (Haxby et al., 2001), and are involved in visual memory encoding (Stern et al., 1996). The right fusiform gyrus is implicated in category-independent recognition (Pourtois et al., 2009) and feature encoding (Garoff et al., 2005). Several areas within the prefrontal cortex are involved in episodic memory (Rugg et al., 2002). Wagner et al. (1998) showed that the ability to recognize verbal material correlated with the magnitude of activation in the left prefrontal cortex, including the inferior frontal gyrus and operculum, during encoding, suggesting a joint role for prefrontal and temporal regions in verbal memory formation. Using positron emission tomography (PET), Frey and Petrides (2002) reported a direct involvement of the right orbitofrontal cortex during encoding, supporting our findings of a relative thickening in this area. The orbitofrontal cortex have strong reciprocal connections with the medial temporal lobe (MTL), known to be heavily involved in memory processing (Simons and Spiers, 2003).

The neurobiological mechanisms responsible for the observed thickness changes are not entirely clear. Use-dependent plasticity of synaptic strength and structure are fundamental mechanisms in memory encoding (Maviel et al., 2004; Trachtenberg et al., 2002). Trachtenberg et al. (2002) have pointed to fast-adjusting neuronal systems, including spine and synapse-turnover. Furthermore, a host of other changes including increase in the size of the soma and nucleus of neurons, glia and capillary dimensions have been demonstrated to influence cortical morphology in animals exposed to enriched

environments (Muotri and Gage, 2006). A recent longitudinal intervention study reported regional grey matter increases in areas showing task-specific activation and concluded that practice-related structural GM changes may primarily be related to synaptic remodeling within specific processing networks (Ilg et al., 2008).

Changes in brain structure in response to 8 weeks of behavioral intervention are consistent with earlier MRI findings reporting GM changes as early as 5 days after intervention (Driemeyer et al., 2008; May et al., 2007). Recent evidence suggests that older adults exhibit intact structural neuroplasticity detectable by MRI (Boyke et al., 2008). In accordance with this, similar structural GM alterations have been reported in old (Boyke et al., 2008) and young adults (Draganski et al., 2004) after juggling practice. Changes in cortical thickness in response to cognitive training have recently been demonstrated in adolescent girls (Haier et al., 2009). The present study is the first to show such training-related changes in an older sample.

Structural correlates of non-experimental memory training have been studied by comparing exceptional memorizers, known to naturally employ some form of the MoL, with controls (Maguire et al., 2002). No structural differences were found between groups, in contrast to the findings in our study. The present study is different in several important aspects. First, and most importantly, we studied an older sample, where none of the participants had any knowledge of the MoL prior to participation. As suggested by Driemeyer et al. (2008), the novelty of the training task may be more critical inducing structural changes than continued training of an already-learned task. Second, our study employs a sensitive data processing scheme utilizing analysis methods specifically designed to detect longitudinal effects on cortical thickness at a submillimeter scale.

An inherent weakness of the present study design is the lack of a priori hypotheses regarding where training-related structural changes might occur. Current knowledge about brain regions involved in MoL training comes from functional imaging studies (Kondo et al., 2005; Maguire et al., 2002; Nyberg et al., 2003) mainly conducted with younger adults. However, Nyberg et al. (2003) studied both older and younger adults utilizing MoL. Some brain regions reported to be engaged while employing MoL (Kondo et al., 2005; Nyberg et al., 2003) show consistency with areas of cortical thickening in our study, e.g. the fusiform gyrus. However, the pattern of cortical thickening is also distributed in other regions (e.g. the right anterior insula) than those merely indicated by functional imaging. The induced structural changes of the current 8-week group-based memory training regimen may not be directly comparable to the functional alterations reported in the previous functional imaging experiments in which the training interventions have lasted less than a day. Thus, one possible explanation of these differences is the large discrepancies in the duration of the interventions between studies. Another, perhaps even more important issue is that associations between functional and structural imaging indices of age-related decline and cognition are far from simple (Persson et al., 2006). As recently discussed by Haier et al (2009), cortical thickness alterations in one area are not necessarily associated with regionally overlapping functional changes as indicated by blood-oxygenation-level dependent (BOLD) contrast in the same area. Further studies on the relationships between the different modalities are needed.

Ageing is associated with cortical thinning (Fjell et al., 2009b; Salat et al., 2004, 2009; Westlye et al., 2009, in press-b). In line with recent longitudinal evidence of 1 year cortical atrophy in healthy adults (Fjell et al., 2009a), cortical thinning was found in nine separate regions in the control group after 8 weeks. Areas of thinning included the four regions reported in the interaction analysis, even though none of the effects survived corrections for multiple comparisons. Importantly, regional thinning was replicated in three of these regions (L1, R1, R2) in an independent replication sample. The right fusiform area (R3) showed no thinning in the replication sample. In addition to real differences in the structural trajectories, this discrepancy might be due to a mean age difference between groups (53.4 vs. 60.3 years) and

possible scanning-related confounds (MR scanner idle time, scanner software updates, and different scan intervals). Further, to validate the current findings we compared the estimated annual whole-brain atrophy rate for our control group (-1.51%) with annualized rates from eight previously published longitudinal studies reviewed in Fotenos et al (2005). Our results are about halfway between these estimates, ranging from -0.37% to -2.1% .

We found an interesting trend relation between TBV and group, indicating some importance of variability in global brain volume over time. No such tendencies were found for the global cortical measures TCV or MCT. Since neither of these relationships were significant we chose to focus on regional differences. The lack of significant global effects of memory training on brain structure supports the use of more regionally fine-grained analysis tools in the study of subtle cortical changes. Previous studies have shown high reliability in thickness estimates over short intervals using FreeSurfer (Han et al., 2006). Thus, we believe the processing and analyses methods employed in the present study are sensitive to short-term changes in regional cortical thickness. Although we attribute the observed thinning to real age-related cerebral changes, it should be emphasized that the magnitude and stability of the short-term cortical thinning should be interpreted with caution, and further large-sample studies on short-term effects of cognitive interventions and age are needed.

Another limitation to the interpretation of the results is the short interval between scans, and follow-ups are needed to establish long-term cognitive and brain structural effects of the training regimen. Further, the stability of the present results should be tested in an independent sample. The histograms in Fig. 5 show the global p -value distribution for both hemispheres, indicating a larger number of vertices showing relative thickening in trainers compared with controls. The fact that all changes in the intervention group were in a positive direction, i.e. regional cortical thickening, whereas this was not the case in the control group, indicates a positive effect of the intervention. In addition to the Monte Carlo simulation and split-half validation, this suggests that it is less likely that the effect in the training group is a result of false-positives. Still, the high costs involved in administering memory training programs limit the number of participants that can be included in such studies. Ideally, larger samples of participants should be studied. In addition to sample size, another weakness of the present study design is the lack of an active control condition. As discussed by (Ball et al., 2002), the use of appropriate control groups has been an understudied issue in earlier cognitive training studies. This also includes the present study and a non-memory control intervention should be included in follow-up studies to further enhance experimental control.

An inherent limitation shared with most imaging studies is that the neurobiological underpinnings of the morphometric changes are unclear. As emphasized by May and Gaser (2006) it is currently unknown what structural plasticity as measured with MRI is driven by at the cellular level and further comparative studies using histology and MRI are needed. Also, recent evidence of short-term micro-structural white matter changes in response to visuo-motor coordination training (Scholz et al., 2009) suggests that future studies should incorporate a multimodal imaging approach to establish converging evidence across structural indices and tissue classes.

Conclusively, our findings indicate that it is possible to influence both specific cognitive performance and short-term trajectory of cortical structure through intensive mental exercise in elderly. A persistent challenge in cognitive aging research is to illuminate and track the various factors contributing to the fact that some age gracefully whereas other decline more rapidly (Buckner, 2004). Our findings of regionally selective longitudinal cortical alterations in response to intensive memory training support that brain change mechanisms are influenced by mental exercise on a short time scale. The results may contribute to an understanding of the cerebral basis for the behavioral findings indicating a role for mental exercise in plasticity and cognitive aging.

Acknowledgments

The Norwegian Research Council (177404/W50) to K.B.W., (175066/D15) to A.M.F. The Medical student research program to A. E. We thank the participants, and Håkon Grydeland and Inge Amlien for help with data acquisition.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2010.05.041.

References

- Anderson, L.A., McConnell, S.R., 2007. The healthy brain and our aging population: Translating science to public health practice. *Alzheimers Dement* 3, S1–S2.
- Ball, K., Berch, D.B., Helmers, K.F., Jobe, J.B., Leveck, M.D., Marsiske, M., Morris, J.N., Rebok, G.W., Smith, D.M., Tennstedt, S.L., Unverzagt, F.W., Willis, S.L., for the ACTIVE Study Group, 2002. Effects of cognitive training interventions with older adults: a randomized controlled trial. *JAMA* 288, 2271–2281.
- 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, 246–254.
- Bower, G.H., 1970. Analysis of a mnemonic device. *Am Sci* 58, 496–510.
- Boyke, J., Driemeyer, J., Gaser, C., Büchel, C., May, A., 2008. Training-induced brain structure changes in the elderly. *J Neurosci* 28, 7031–7035.
- Buckner, R.L., 2004. Memory and executive function in aging and AD: multiple factors that cause decline and reserve factors that compensate. *Neuron* 44, 195–208.
- Buonomano, D.V., Merzenich, M.M., 1998. CORTICAL PLASTICITY: from synapses to maps. *Annu Rev Neurosci* 21, 149–186.
- Corballis, P.M., 2003. Visuospatial processing and the right-hemisphere interpreter. *Brain Cogn* 53, 171–176.
- Craig, A.D., 2009. How do you feel—now? The anterior insula and human awareness. *Nat Rev Neurosci* 10, 59–70.
- Dale, A.M., Sereno, M.I., 1993. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. *J Cogn Neurosci* 5, 162–176.
- Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 9, 179–194.
- DeFelipe, J., 2006. Brain plasticity and mental processes: Cajal again. *Nat Rev Neurosci* 7, 811–817.
- Delis, D.C., Kramer, J.H., Kaplan, E., Ober, B.A., 2000. California Verbal Learning Test—Second Edition (CVLT-II). The Psychological Corporation, San Antonio, TX.
- Dosenbach, N.U.F., Fair, D.A., Miezin, F.M., Cohen, A.L., Wenger, K.K., Dosenbach, R.A.T., Fox, M.D., Snyder, A.Z., Vincent, J.L., Raichle, M.E., Schlaggar, B.L., Petersen, S.E., 2007. Distinct brain networks for adaptive and stable task control in humans. *Proc Natl Acad Sci* 104, 11073–11078.
- Draganski, B., May, A., 2008. Training-induced structural changes in the adult human brain. *Behav Brain Res* 192, 137–142.
- Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., May, A., 2004. Neuroplasticity: changes in grey matter induced by training. *Nature* 427, 311–312.
- Draganski, B., Gaser, C., Kempermann, G., Kuhn, H.G., Winkler, J., Büchel, C., May, A., 2006. Temporal and spatial dynamics of brain structure changes during extensive learning. *J Neurosci* 26, 6314–6317.
- Driemeyer, J., Boyke, J., Gaser, C., Büchel, C., May, A., 2008. Changes in gray matter induced by learning—revisited. *PLoS ONE* 3.
- 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, 11050–11055.
- Fischl, B., Sereno, M.I., Dale, A.M., 1999a. Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9, 195–207.
- Fischl, B., Sereno, M.I., Tootell, R.B., Dale, A.M., 1999b. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp* 8, 272–284.
- 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.
- Fjell, A.M., Walhovd, K.B., Fennema-Notestine, C., McEvoy, L.K., Hagler, D.J., Holland, D., Brewer, J.B., Dale, A.M., 2009a. One-year brain atrophy evident in healthy aging. *J Neurosci* 29, 15223–15231.
- Fjell, A.M., Westlye, L.T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., Agartz, I., Salat, D.H., Greve, D.N., Fischl, B., Dale, A.M., Walhovd, K.B., 2009b. High consistency of regional cortical thinning in aging across multiple samples. *Cereb Cortex* 19, 2001–2012.
- 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.
- Forman, S., Cohen, J., Fitzgerald, M., Eddy, W., Mintun, M., Noll, D., 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med* 33, 636–647.
- 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, 1032–1039.

- Frey, S., Petrides, M., 2002. Orbitofrontal cortex and memory formation. *Neuron* 36, 171–176.
- Garoff, R.J., Slotnick, S.D., Schacter, D.L., 2005. The neural origins of specific and general memory: the role of the fusiform cortex. *Neuropsychologia* 43, 847–859.
- Hagler Jr., D.J., Saygin, A.P., Sereno, M.I., 2006. Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. *Neuroimage* 33, 1093–1103.
- Haier, R., Karama, S., Leyba, L., Jung, R., 2009. MRI assessment of cortical thickness and functional activity changes in adolescent girls following three months of practice on a visual-spatial task. *BMC Res Notes* 2, 174.
- 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.
- Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430.
- Hayasaka, S., Nichols, T.E., 2003. Validating cluster size inference: random field and permutation methods. *Neuroimage* 20, 2343–2356.
- Ilg, R., Wohlschlagel, A.M., Gaser, C., Liebau, Y., Dauner, R., Woller, A., Zimmer, C., Zihl, J., Muhlau, M., 2008. Gray matter increase induced by practice correlates with task-specific activation: a combined functional and morphometric magnetic resonance imaging study. *J Neurosci* 28, 4210–4215.
- Jennings, J.M., Webster, L.M., Kleykamp, B.A., Dagenbach, D., 2005. Recollection training and transfer effects in older adults: successful use of a repetition-lag procedure. *Aging Neuropsychol Cogn* 12, 278–298.
- Jernigan, T.L., Archibald, S.L., Fennema-Notestine, C., Gamst, A.C., Stout, J.C., Bonner, J., Hesselink, J.R., 2001. Effects of age on tissues and regions of the cerebrum and cerebellum. *Neurobiol Aging* 22, 581–594.
- Johansson, B.B., 2004. Brain plasticity in health and disease. *Keio J Med* 53, 231–246.
- Jones, S., Nyberg, L., Sandblom, J., Stigsdotter Neely, A., Ingvar, M., Magnus Petersson, K., Backman, L., 2006. Cognitive and neural plasticity in aging: general and task-specific limitations. *Neurosci Biobehav Rev* 30, 864–871.
- 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.
- Kliegl, R., Smith, J., Baltes, P.B., 1990. On the locus and process of magnification of age differences during mnemonic training. *Dev Psychol* 26, 894–904.
- Kondo, Y., Suzuki, M., Mugikura, S., Abe, N., Takahashi, S., Iijima, T., Fujii, T., 2005. Changes in brain activation associated with use of a memory strategy: a functional MRI study. *Neuroimage* 24, 1154–1163.
- Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., Baker, C.I., 2009. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 12, 535–540.
- Kuperberg, G.R., Broome, M.R., McGuire, P.K., David, A.S., Eddy, M., Ozawa, F., Goff, D., West, W.C., Williams, S.C.R., van der Kouwe, A.J.W., Salat, D.H., Dale, A.M., Fischl, B., 2003. Regionally localized thinning of the cerebral cortex in schizophrenia. *Arch Gen Psychiatry* 60, 878–888.
- Maguire, E.A., Valentine, E.R., Wilding, J.M., Kapur, N., 2002. Routes to remembering: the brains behind superior memory. *Nat Neurosci* 90–95.
- 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*. Elsevier, pp. 81–109.
- Martin, S.J., Grimwood, P.D., Morris, R.G.M., 2000. Synaptic plasticity and memory: an evaluation of the hypothesis. *Annu Rev Neurosci* 23, 649–711.
- Maviel, T., Durkin, T.P., Menzaghi, F., Bontempi, B., 2004. Sites of neocortical reorganization critical for remote spatial memory. *Science* 305, 96–99.
- May, A., Gaser, C., 2006. Magnetic resonance-based morphometry: a window into structural plasticity of the brain. *Curr Opin Neurol* 19, 407–411.
- May, A., Hajak, G., Ganssbauer, S., Steffens, T., Langguth, B., Kleinjung, T., Eichhammer, P., 2007. Structural brain alterations following 5 days of intervention: dynamic aspects of neuroplasticity. *Cereb Cortex* 17, 205–210.
- Muotri, A.R., Gage, F.H., 2006. Generation of neuronal variability and complexity. *Nature* 441, 1087–1093.
- National Council on the Aging, 2000. American perceptions of aging in the 21st century. National Council on the Aging, Washington.
- Nyberg, L., Sandblom, J., Jones, S., Neely, A.S., Petersson, K.M., Ingvar, M., Backman, L., 2003. Neural correlates of training-related memory improvement in adulthood and aging. *Proc Natl Acad Sci U S A* 100, 13728–13733.
- Pascual-Leone, A., Amedi, A., Fregni, F., Merabet, L.B., 2005. The plastic human brain cortex. *Annu Rev Neurosci* 28, 377–401.
- Persson, J., Nyberg, L., Lind, J., Larsson, A., Nilsson, L.-G., Ingvar, M., Buckner, R.L., 2006. Structure–function correlates of cognitive decline in aging. *Cereb Cortex* 16, 907–915.
- Posner, M.I., Petersen, S.E., 1990. The attention system of the human brain. *Annu Rev Neurosci* 13, 25–42.
- Pourtois, G., Schwartz, S., Spiridon, M., Martuzzi, R., Vuilleumier, P., 2009. Object representations for multiple visual categories overlap in lateral occipital and medial fusiform cortex. *Cereb Cortex* 19, 1806–1819.
- Raz, N., Rodrigue, K.M., 2006. Differential aging of the brain: patterns, cognitive correlates and modifiers. *Methodological and conceptual advances in the study of brain-behavior dynamics: a multivariate lifespan perspective*, 30, p. 730.
- 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, 433–438.
- Raz, N., Rodrigue, K.M., Haacke, E.M., 2007. Brain aging and its modifiers: insights from in vivo neuromorphometry and susceptibility weighted imaging. *Ann NY Acad Sci* 1097, 84–93.
- 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, 53–61.
- Reid, L., MacLulich, A., 2006. Subjective memory complaints and cognitive impairment in older people. *Dement Geriatr Cogn Disord* 22, 471–485.
- Rosas, H.D., Liu, A.K., Hersch, S., Glessner, M., Ferrante, R.J., Salat, D.H., van der Kouwe, A., Jenkins, B.G., Dale, A.M., Fischl, B., 2002. Regional and progressive thinning of the cortical ribbon in Huntington's disease. *Neurology* 58, 695–701.
- Rugg, M.D., Otten, L.J., Henson, R.N., 2002. The neural basis of episodic memory: evidence from functional neuroimaging. *Philos Trans R Soc Lond B Biol Sci* 357, 1097–1110.
- Salat, D.H., Buckner, R.L., Snyder, A.Z., Greve, D.N., Desikan, R.S., Busa, E., Morris, J.C., Dale, A.M., Fischl, B., 2004. Thinning of the cerebral cortex in aging. *Cereb Cortex* 14, 721–730.
- Salat, D.H., Lee, S.Y., van der Kouwe, A.J., Greve, D.N., Fischl, B., Rosas, H.D., 2009. Age-associated alterations in cortical gray matter and subjacent white matter signal intensity. *Neuroimage* 47, S113.
- Scholz, J., Klein, M.C., Behrens, T.E.J., Johansen-Berg, H., 2009. Training induces changes in white-matter architecture. *Nat Neurosci* 12, 1370–1371.
- Simons, J.S., Spiers, H.J., 2003. Prefrontal and medial temporal lobe interactions in long-term memory. *Nat Rev Neurosci* 4, 637–648.
- Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging* 17, 87–97.
- Sridharan, D., Levitin, D.J., Menon, V., 2008. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proc Natl Acad Sci* 105, 12569–12574.
- Stern, C.E., Corkin, S., Gonzalez, R.G., Guimaraes, A.R., Baker, J.R., Jennings, P.J., Carr, C.A., Sugiura, R.M., Vethanatham, V., Rosen, B.R., 1996. The hippocampal formation participates in novel picture encoding: evidence from functional magnetic resonance imaging. *Proc Natl Acad Sci USA* 93, 8660–8665.
- Trachtenberg, J.T., Chen, B.E., Knott, G.W., Feng, G., Sanes, J.R., Welker, E., Svoboda, K., 2002. Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. *Nature* 420, 788–794.
- Valenzuela, M.J., Sachdev, P., 2006a. Brain reserve and cognitive decline: a non-parametric review. *Psychol Med* 36, 1065–1073.
- Valenzuela, M.J., Sachdev, P., 2006b. Brain reserve and dementia: a systematic review. *Psychol Med* 36, 441–454.
- Valenzuela, M.J., Jones, M., Wen, W., Rae, C., Graham, S., Shnier, R., Sachdev, P., 2003. Memory training alters hippocampal neurochemistry in healthy elderly. *NeuroReport* 14, 1333–1337.
- van der Kouwe, A.J., Benner, T., Fischl, B., Schmitt, F., Salat, D.H., Harder, M., Sorensen, A.G., Dale, A.M., 2005. On-line automatic slice positioning for brain MR imaging. *Neuroimage* 27, 222–230.
- Verhaeghen, P., Marcoen, A., Goossens, L., 1992. Improving memory performance in the aged through mnemonic training: a meta-analytic study. *Psychol Aging* 7, 242–251.
- Vul, E., Harris, C., Winkelman, P., Pashler, H., 2009. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspect Psychol Sci* 4, 274–290.
- Wagner, A.D., Schacter, D.L., Rotte, M., Koutstaal, W., Maril, A., Dale, A.M., Rosen, B.R., Buckner, R.L., 1998. Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science* 281, 1188–1191.
- Walhovd, K.B., Fjell, A.M., Reinvang, I., Lundervold, A., Dale, A.M., Eilertsen, D.E., Quinn, B.T., Salat, D., Makris, N., Fischl, B., 2005. Effects of age on volumes of cortex, white matter and subcortical structures. *Neurobiol Aging* 26, 1261–1270 discussion 1275–1268.
- 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., 2009. Consistent neuroanatomical age-related volume differences across multiple samples. *Neurobiology of Aging* in press, Corrected Proof.
- Ward, B.D., 2000. Simultaneous Inference for fMRI Data. AFNI 3dDeconvolve Documentation, Medical College of Wisconsin.
- Wechsler, D., 1999. Wechsler abbreviated scale of intelligence. The Psychological Corporation, San Antonio, TX.
- Westlye, L.T., Walhovd, K.B., Dale, A.M., Espeseth, T., Reinvang, I., Raz, N., Agartz, I., Greve, D.N., Fischl, B., Fjell, A.M., 2009. Increased sensitivity to effects of normal aging and Alzheimer's disease on cortical thickness by adjustment for local variability in gray/white contrast: a multi-sample MRI study. *Neuroimage* 47, 1545–1557.
- Westlye, L.T., Walhovd, K.B., Dale, A.M., Bjørnerud, A., Due-Tønnessen, P., Engvig, A., Grydeland, H., Tamnes, C.K., Østby, Y., Fjell, A.M., in press-a. Life-Span Changes of the Human Brain White Matter: Diffusion Tensor Imaging (DTI) and Volumetry. *Cereb. Cortex*.
- Westlye, L.T., Walhovd, K.B., Dale, A.M., Bjørnerud, A., Due-Tønnessen, P., Engvig, A., Grydeland, H., Tamnes, C.K., Østby, Y., Fjell, A.M., in press-b. Differentiating maturational and aging-related changes of the cerebral cortex by use of thickness and signal intensity. *Neuroimage*.
- Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M., Leirer, V.O., 1983. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 17, 37–49.