◆ Human Brain Mapping 33:2390–2406 (2012) ◆

Memory Training Impacts Short-Term Changes in Aging White Matter: A Longitudinal Diffusion Tensor Imaging Study

Andreas Engvig,[1]* Anders M. Fjell,[1,2] Lars T. Westlye,[1] Torgeir Moberget,[1] Øyvind Sundseth,[2] Vivi Agnete Larsen,[1] and Kristine B. Walhovd[1,2]

[1]*Center for the Study of Human Cognition, Department of Psychology, University of Oslo, Oslo, Norway*
[2]*Department of Neuropsychology, Oslo University Hospital, Ullevaal, Norway*

Abstract: A growing body of research indicates benefits of cognitive training in older adults, but the neuronal mechanisms underlying the effect of cognitive intervention remains largely unexplored. Neuroimaging methods are sensitive to subtle changes in brain structure and show potential for enhancing our understanding of both aging- and training-related neuronal plasticity. Specifically, studies using diffusion tensor imaging (DTI) suggest substantial changes in white matter (WM) in aging, but it is not known whether cognitive training might modulate these structural alterations. We used tract-based spatial statistics (TBSS) optimized for longitudinal analysis to delineate the effects of 8 weeks intensive memory training on WM microstructure. 41 participants (mean age 61 years) matched for age, sex and education were randomly assigned to an intervention or control group. All participants underwent MRI-scanning and neuropsychological assessments at the beginning and end of the study. Longitudinal analysis across groups revealed significant increase in frontal mean diffusivity (MD), indicating that DTI is sensitive to WM structural alterations over a 10-week interval. Further, group analysis demonstrated positive effects of training on the short-term changes. Participants in the training group showed a relative increase in fractional anisotropy (FA) compared with controls. Further, a significant relationship between memory improvement and change in FA was found, suggesting a possible functional significance of the reported changes. The training effect on FA seemed to be driven by a relative decrease in radial diffusivity, which might indicate a role for myelin-related processes in WM plasticity. *Hum Brain Mapp 33:2390–2406, 2012*.

Key words: plasticity; white matter; memory; training; aging; DTI; FA; longitudinal

Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: The Research Council of Norway; Contract grant number: 177404/W50, 175066/D15; Contract grant sponsor: Kavli Research Center for Ageing and Dementia.

*Correspondence to: Andreas Engvig, Center for the Study of Human Cognition, Department of Psychology, University of Oslo, Oslo 0317, Norway. E-mail: andreas.engvig@gmail.com

Received for publication 19 November 2010; Revised 17 March 2011; Accepted 5 May 2011

DOI: 10.1002/hbm.21370

Published online 5 August 2011 in Wiley Online Library (wileyonlinelibrary.com).

INTRODUCTION

Numerous studies show behavioral benefits of cognitive training in both healthy and pathological aging [see Buschert et al., 2010; Reichman et al., 2010, for recent reviews]. However, the neuronal substrates of the observed behavioral improvements are largely unknown. Neuroimaging is a promising tool for elucidating the underlying neurobiological mechanisms of training [Lustig et al., 2009]. Recent studies of cognitive interventions in older adults have pointed to prefrontal functional [Mozolic et al., 2010] and structural [Engvig et al., 2010] reorganization of the cortical gray matter (GM). Studies using other imaging modalities might enhance our understanding of these observed changes.

One promising modality is diffusion tensor imaging (DTI) enabling characterization of white matter (WM) microstructure in vivo [Le Bihan, 2003]. DTI-indices of WM show considerable age-related differences throughout life [Burzynska et al., 2010; Giorgio et al., 2010; Madden et al., 2004; Pfefferbaum and Sullivan, 2003; Pfefferbaum et al., 2005]. Putative DTI markers of WM aging include increasing mean diffusivity (MD) and decreasing fractional anisotropy (FA) [see Chanraud et al., 2010, for a review].

Life-span studies have shown that FA decreases and MD increases are apparent in most regions by the end of the third decade of life [Kennedy and Raz, 2009b]. In a study of 430 subjects aged 8–85 years, estimated age of maximum FA was 29.1 years for whole-brain WM [Westlye et al., 2010]. Regional analyses have shown an anterior-posterior gradient [Davis et al., 2009; Salat et al., 2005], with stronger age effects observed in anterior as compared to posterior regions [Bennett et al., 2010; Head et al., 2004].

The observed WM differences are likely related to aging of specific cognitive skills [see Johansen-Berg, 2010, for a review]. Kennedy and Raz [2009a] reported that age-related WM degradation, as indexed by FA and MD, was related to multiple measures of cognitive performance. Moreover, some studies indicate that learning cognitive skills is reflected in FA variability. For instance, Flöel et al. [2009] demonstrated that higher FA scores in WM adjacent to Broca's area predicted grammar-learning success. DTI-studies of expertise have shown that training-related differences in FA [Hanggi et al., 2010], and MD [Imfeld et al., 2009] correlates with amount of time spent practicing, pointing to experience-related WM reorganization traceable in vivo [see Jäncke, 2009, for a review].

Recently, longitudinal studies of younger adults have confirmed experience-related WM changes over short intervals after motor [Scholz et al., 2009] and cognitive [Takeuchi et al., 2010] training, respectively. Keller and Just [2009] showed that remediation reading training in children is followed by significant FA increases associated with successful training. Finally, Lövdén et al. [2010] reported that age-related WM changes in the anterior corpus callosum were influenced by an intensive 100-days cognitive intervention scheme, indicating intact WM plasticity in elderly. The four longitudinal studies reviewed above consistently reported FA-increases following intervention. The study by Lövdén also reported training-related decreases in MD. Interestingly, these patterns contrast the observed age-related trajectories described above, with decreasing FA and increasing MD.

Age-related changes in WM diffusivity are likely to reflect increase in brain water content, demyelination, disruption of axon structure, and overall rarefaction of nerve fibers [for a review, see Minati et al., 2007]. In one study, age-related decreases in FA were mainly driven by increases in radial diffusivity (RD) [Giorgio et al., 2010]. RD, together with FA and MD, was shown to be a robust predictor of myelin content in post-mortem human brains [Schmierer et al., 2008]. Increased RD has also been related to myelin damage in animal models of multiple sclerosis [Sun et al., 2007] and dysmyelination [Song et al., 2002]. On the contrary, increased CNS myelination has been reported after neuronal activity [Ishibashi et al., 2006; Stevens et al., 2002] and is suggested to play a role in learning and cognition [see Fields, 2008, for a review]. Inferences such as linking RD and myelination are, however, currently limited by the challenges of interpreting changes in diffusion MRI measures in biological terms [Beaulieu, 2009].

Despite the limitations of current DTI-interpretation, initial studies reviewed above have demonstrated potential for experience-related cerebral WM reorganization in humans. Studies of training-related WM effects in older individuals are, nevertheless, rare. Besides the need for replication of such studies, unanswered questions include the temporal aspects of WM changes in response to training and the relationship between WM and behavioral change. A more precise understanding of plasticity-related mechanisms in aging is important because of the potentially broad implications for guiding development of intervention programs aimed at halting or reversing age-related cognitive decline and neurodegeneration.

In the present study we assessed short-term WM changes using a tract-based spatial statistics (TBSS) [Smith et al., 2006] protocol optimized for longitudinal analysis. There were three main aims:

First, we tested whether changes in DTI-indices of WM were detectable after 10-weeks in 41 middle-aged and older individuals. As reviewed above, a highly consistent finding in previous studies is FA decreases and MD increases distributed with an anterior-posterior gradient. We therefore hypothesized that if changes were detectable; they were likely to be manifested as frontally distributed decreases in FA and/or increases in MD.

Second, we investigated the hypothesis of WM plasticity in response to memory intervention by randomly assigning the participants to either an intervention group, participating in 8 weeks of intensive verbal memory training or a control group. On the basis of the previous studies of cognitive training effects on cerebral WM [Lövdén et al., 2010], we hypothesized group × time interactions in

regions showing age-related changes, effectively demonstrating that the training regimen influenced the rate of short-term WM deterioration.

The two first hypotheses were tested voxel-wise for FA and MD. To further examine the hypothesis of training-related effects, follow-up analyses were performed for RD and AD within regions showing group × time interactions. This approach reduces the number of statistical comparisons, and allows for more specific interpretation of changes in FA and MD as opposed to considering the different metrics in four independent voxel-wise analyses. Since evidence reviewed above suggests a possible role for myelination in both aging and cognition, we hypothesized that training would primarily impact RD changes in these regions.

Third, on the basis of our previous finding of a correlation between frontal cortical thickness changes and training-related memory improvements [Engvig et al., 2010], we hypothesized that the intervention-related WM effects would be largest for the subjects showing the most beneficial effects on memory performance.

MATERIALS AND METHODS

To assess the effects of cognitive intervention on age-related trajectories in WM, we used a randomized controlled design in the present study. The study sample is the same as in a previous article focusing on cortical thickness [Engvig et al., 2010], and a comprehensive description of the intervention and behavioral improvements are reported there.

Sample

Forty eight volunteers were recruited through a local newspaper ad and assessed for eligibility and screened using a structured health interview before inclusion. All included participants were required to be right-handed ("are you right-handed?" yes/no-dichotomy, confirmed for hand writing and drawing by observation) 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.

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). 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] and IQ < 85, derived from the vocabulary and matrices sub tests of the Wechsler Abbreviated Scale of Intelligence (WASI) [Wechsler, 1999]. One participant filled out the GDS form after completion of the program, but 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 two standard deviations (SD) or more below population norm were also excluded. One participant was excluded on the basis of the CVLT score. Table I presents demographic and neuropsychological measures for each group at baseline. Briefly, there were no significant baseline differences between groups in any of the demographic or neuropsychological variables. Mean age and sex-adjusted *T*-scores for CVLT learning and long delay was 59.3 (9.0) and 57.7 (8.7), respectively. This indicates that the sample mean was about 1 SD above population norms with regards to both IQ (see Table I) and verbal memory functioning.

TABLE I. Characteristics of completing participants at baseline (*n* = 41)

	Intervention (11F/10M)		Control (11F/9M)		
	M	SD	M	SD	*P*-value*
Age	61.7	9.4	60.3	9.1	0.64
Education	15.1	1.9	15.6	1.8	0.45
IQ	117.8	9.0	118.8	9.2	0.74
MMSE	29.0	1.0	29.1	0.9	0.63
GDS	1.7	1.9	1.5	2.2	0.75
TMT A	37.1	16.1	33.0	13.8	0.39
TMT B	81.6	35.3	70.7	32.7	0.31
Digit-symbol	65.2	15.6	67.7	14.1	0.61
Rey-O, recall	19.5	6.5	21.1	5.7	0.42
CVLT, 1–5 total	50.4	9.1	52.3	10.7	0.54
CVLT, long delay	11.8	2.3	11.9	3.6	0.93
Re-test interval	65.4	6.8	65.3	9.5	0.95

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 min delayed recall of the Rey-Osterreith complex figure test; CVLT 1-5 total, score with intrusions subtracted; CVLT long delay, score on 20 min delay recall with intrusions subtracted. Re-test interval denotes days between 1st and 2nd MRI scanning session.
**P*-values are two-tailed and based on independent samples *t* test with intervention or control as grouping variable.

Design

At the start of the project - time point one (tp1)—all participants were tested on a series of neuropsychological and behavioral tests, and underwent an MRI scan. Next, participants were randomly assigned to either of two groups: (1)

an intervention group (ages 42–76 years, $n = 21$) participating in an 8-week intensive verbal memory training program, or (2) a control group (ages 42–77 years, $n = 20$). The groups were carefully matched by sex, age, and education level. The following week, the intervention group started the training scheme, while the controls were instructed to continue living as usual. At time point two (tp2), the week after completion of the intervention, all participants were reexamined on selected behavioral tests and underwent a second MRI scan. Because of practical and logistic reasons, all individuals were reassessed in the week after the end of the intervention. Mean interval between the two scanning sessions was 65.3 days (SD = 8.1). There were no differences in interval between groups ($t = 0.07$, $P = 0.95$). After group allocation, one participant in the training group was excluded due to inadequate MRI data quality, and one was excluded due to missing MRI data. Two participants in the control group withdrew consent. Thus, 41 participants (21 training; 20 control) were included in the final analysis.

Intervention

For the present study we developed a memory training program aimed at improving serial verbal recollection memory by implementing the mnemonic technique Method of loci (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 classroom session and 4 days of homework exercises. The classroom 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. In total, participants finished 40 exercises (in-class + homework), comprising roughly 25 min of training 5-days a week for 8 weeks. An extensive description of the intervention, including examples of in-class and homework exercises used in the memory training program has been published in [Engvig et al., 2010].

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 six 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 (as a proxy of 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 judgments/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 Data Acquisition

Imaging data were collected using a 12 channel head coil on the same 1.5 Tesla Siemens Avanto scanner (Siemens Medical Solutions, Erlangen, Germany) at Oslo University Hospital, Rikshospitalet. The same scanner software and software version were used at both time-points. For DTI a single-shot twice-refocused spin echo echo planar imaging (EPI) pulse sequence with 30 diffusion sensitized gradient directions and the following parameters was used: repetition time (TR)/echo time (TE) = 8,200 ms/82 ms, b-value = 700 s mm^{-2}, voxel size = 2.0 × 2.0 × 2.0 mm^3. This sequence is optimized to minimize eddy-current-induced image distortions [Reese et al., 2003]. The sequence was repeated in two successive runs with 10 non-diffusion-weighted images ($b = 0$) in addition to 30 diffusion weighted images collected per acquisition. Each volume

consisted of 64 axial slices. Total scanning time was 11 min, 21 s. All acquisitions were visually inspected for imaging artifacts (e.g., movement, susceptibility artifacts). In addition, a T2-weighted fluid-attenuated inversion recovery (FLAIR) sequence was used to aid the neuroradiological examination. All MRI scans were subjected to a radiological evaluation by an experienced 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.

DTI Analysis

All datasets were processed and analyzed at the Center for the Study of Human Cognition, University of Oslo, with additional use of computing resources from the Titan High Performance Computing facilities (http://hpc.uio.no/index.php/Titan). Image analyses and tensor calculations were done using FSL (http://www.fmrib.ox.ac.uk/fsl/index.html) [Smith et al., 2004]. First, the raw 4D datasets were corrected for distortions due to eddy currents and head motion between volumes by using an affine registration to the first $b = 0$ volume by means of FLIRT [Jenkinson and Smith, 2001]. The b matrix was rotated accordingly to preserve the orientational information in the data. Next, the diffusion tensor was fitted to the data to calculate the FA, eigenvector and eigenvalue maps. MD was defined as the mean of all three eigenvalues $((\lambda_1 + \lambda_2 + \lambda_3)/3)$, radial diffusion (RD) as the mean of the second and third eigenvalues $((\lambda_1 + \lambda_2)/2)$, and axial diffusion (AD) as the principal diffusion eigenvalue (λ_1).

Longitudinal TBSS-Protocol

We performed voxel-wise analyses of FA and MD maps across subjects and time-points using TBSS [Smith et al., 2006], a method aimed to solve issues of cross-subject data alignment. In addition to cross-subject alignment, studies of longitudinal brain change need to account for residual variation due to multiple time-points. Because of possible variations in imaging geometry over time due to gradient calibration drift, local field distortions, or varying head placement in the scanner, subtle changes in brain structure cannot be measured before the images for the two time-points have been properly aligned. Thus, we modified the TBSS default scheme to optimize anatomical longitudinal alignment in the present study.

The longitudinal processing scheme consisted of several steps (see Fig. 1 for a schematic overview). The initial registration involves a robust and accurate automated linear registration using FLIRT [Jenkinson and Smith, 2001] with a special optimization schedule validated for accurate longitudinal studies of brain atrophy [Smith et al., 2007]. For each subject, we used the brain extraction tool (BET) [Smith, 2002] to extract brain and skull images from the $b = 0$ volumes at each time-point. The initial alignment between the two time-points is informed by the extracted skull representation (which is not assumed to change during the course of the intervention) to constrain the registration scaling. Next, both volumes are resampled into the space halfway between the two [Smith et al., 2001], which only requires a single registration per volume and thus minimizes registration bias towards one of the time-points.

Next, we averaged the two halfway registered FA-maps to generate a subject-wise mid-space template (base FA template). Then, TBSS automatically aligned each subject's base FA template to every other one and identified a study-specific target image (the best target of all FA templates in sample). The chosen target image was then affine-aligned into standard space ($1 \times 1 \times 1$ mm^3). Next, all base FA templates were transformed into standard space by combining (1) a non-linear transform to the study-specific target image with (2) the affine transform from the study-specific target image to standard space. We averaged the transformed base FA templates across participants to generate a study-specific mean FA map. The mean FA map was 'thinned' and thresholded at FA > 0.2 to generate a WM tract skeleton representing the center of the tracts common to all subjects. FA values from the base FA template for each individual were warped onto this mean skeleton by searching perpendicular from the center of the skeleton for maximum FA. Next, the same skeletonization procedure was applied to the properly aligned DTI maps from the two time-points after applying a small smoothing kernel (sigma = 2). The statistical analyses performed are inherently voxel-wise, and small registration errors may decrease the sensitivity to regional effects. Therefore, the smoothing step was performed to further account for residual variation between time points.

Statistical Analyses

The voxel-wise DTI analyses were carried out using permutation-based cross-subject statistics [Nichols and Holmes, 2002] implemented in randomise in FSL. 10,000 permutations were performed for each contrast and statistical inference was made using threshold-free cluster enhancement (TFCE) [Smith and Nichols, 2009].

Our first hypothesis was that short-term changes are detectable across groups as decreases in FA and/or increases in MD. We tested for this in a whole-brain voxel-wise analysis by applying appropriate contrasts within the framework of the General Linear Model (GLM). Effects were considered significant at $P < 0.05$, after correction for multiple comparisons across space. To test for age effects on the longitudinal alterations in the DTI measures, we extracted individual mean values from significant clusters generated in the voxel-wise analysis and tested for effects of age using Spearman's nonparametric correlation analysis. Previous literature did not provide clear-cut indications of what to expect in terms of RD and AD changes. We therefore selected the more established MD and

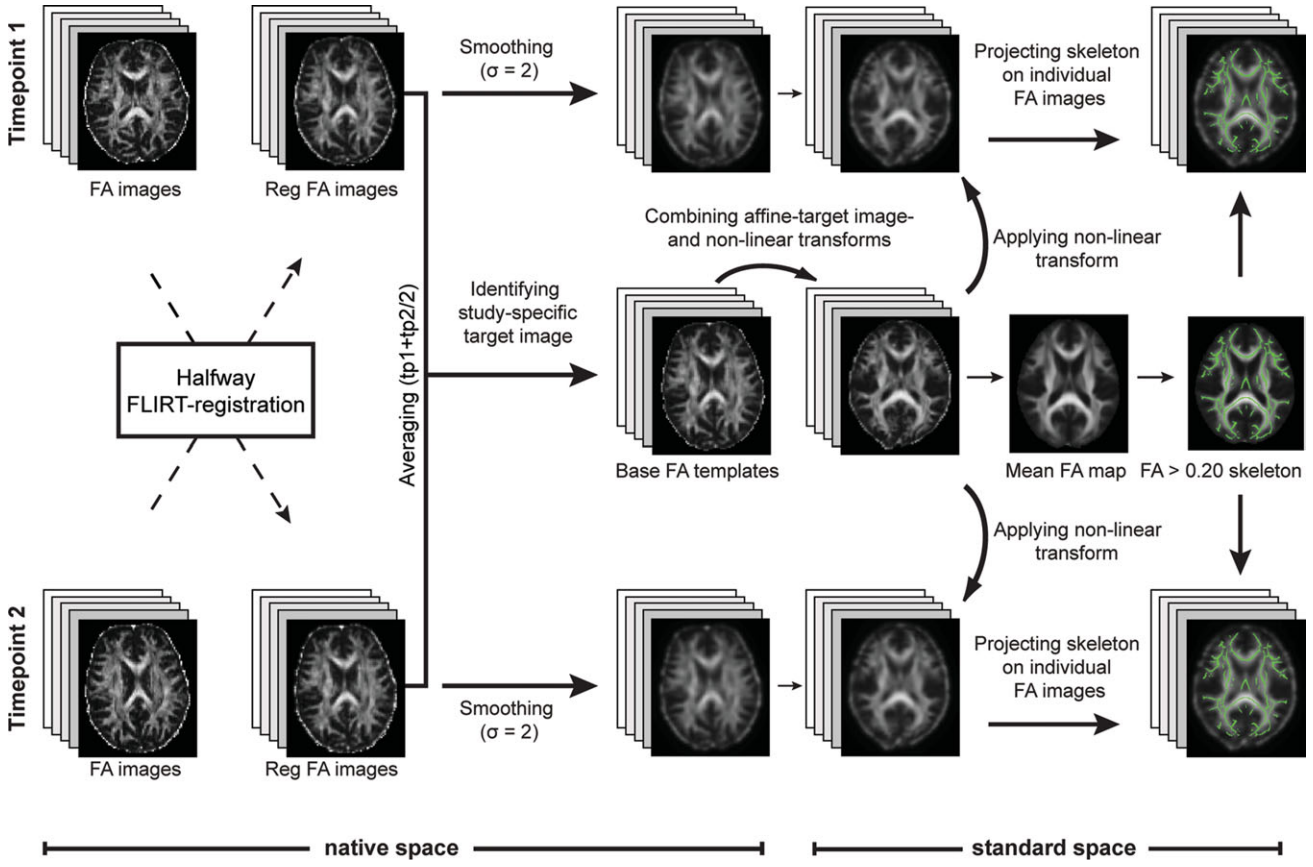

Figure 1.

Flow chart of steps used to process longitudinal DTI data. FA-volumes for each time-point for each subject are linearly registered to a space halfway between tp1 and tp2 using FLIRT and the extracted skull representation. From these halfway registered FA volumes base FA templates are calculated for each subject by averaging the two time-points. TBSS automatically identifies the best (most representative) of these templates and selects this as a target image for transformation of all base FA templates into a standard 1 × 1 × 1 space. The warped FA templates are used to generate a mean FA map of all subjects. The mean FA map is thinned to create the mean FA skeleton, common to all subjects. The skeleton is thresholded and binarized at FA > 0.20 to reduce the likelihood of partial voluming effects. To further account for residual variation, each participant's halfway FA volumes are applied a small smoothing kernel (sigma = 2), to further account for residual variation. After warping individual FA-volumes to standard space, the mean FA skeleton is projected to the individual volumes for each time-point for use in the statistical analyses. The halfway registration step, smoothing, warping and projection are subsequently performed for MD, RD and AD using the mean FA skeleton.

FA-metrics for voxel-wise analyses in the present study. This reduced the number of statistical analyses run.

To assess the notion of an effect-gradient from posterior to anterior segments, we evaluated the distribution of effects from the longitudinal analyses along the *y*-axis. Specifically, mean statistical *t* values for each *y* coordinate within the WM skeleton was fitted by means of robust locally weighted polynomial regression (rLOESS) as a function of MNI *y*-coordinate integer [Cleveland and Devlin, 1988]. The coordinate-wise rLOESS fitting was done using custom Matlab routines. Since rLOESS function does not represent a specific model expression, no formal tests of significance were implemented.

Second, we hypothesized that memory training would impact longitudinal changes in frontal regions. Thus, instead of assessing training-related effect across the entire WM skeleton, we selected frontal regions showing significant longitudinal effects across groups for the following analyses of training-related effects. This reduced the number of statistical comparisons and the likelihood of Type I errors.

To investigate if there was an effect of training, we tested for group × time interaction effects. Specifically, we

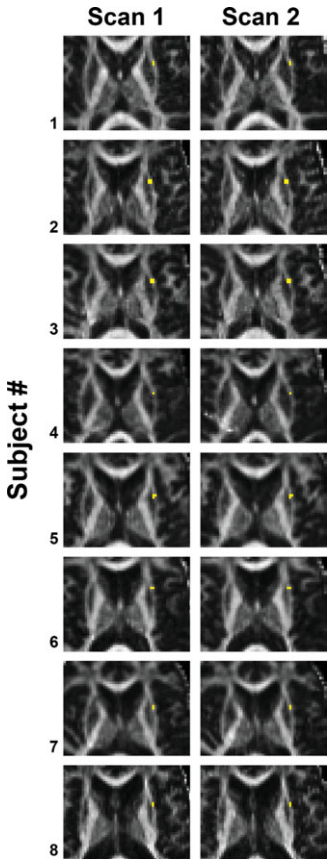

Figure 2.

Anatomical position of de-projected voxels at each tp for eight participants (Horizontal section, radiological convention, *Z*-coordinate (mm) = 20.34). The figure shows a small random cluster in the left anterior external capsule found using a *t*-map from the across-group longitudinal analysis de-projected to individual FA volumes. These example volumes comprise subjects showing the most positive change ($n = 4$, top four) and the least positive change ($n = 4$, bottom four) in diffusivity in this area. The figure shows that the registration and alignment procedure applied in the present study assures excellent anatomical overlap across scans.

tested voxel-wise if group membership influenced changes in FA and MD. To exclude the possibility that baseline differences between groups influenced the longitudinal effects, we first performed a whole-brain analysis comparing MD and FA at tp1 between groups, and also included baseline values as voxel-wise regressors in the statistical models. Significant group × time interaction effects ($P < 0.05$, corrected) were further examined by running separate analyses on mean change in radial (RD) and axial (AD) diffusivity.

Third, we tested for relationships between rate of WM microstructural changes and training-related memory improvement. Specifically, we used Spearman's nonparametric correlation analysis to test for an association between change in FA and improvement in verbal source memory performance.

Evaluation of the Spatial Alignment Between Time-Points

To evaluate the alignment of the DTI data between time-points, we de-projected voxels from the statistical *t*-maps onto each individual's FA map for each time-point. Figure 2 shows an example of the accuracy of the alignment. The figure shows one de-projected cluster (left anterior external capsule) for eight participants for each time-point, indicating excellent spatial correspondence between the two time-points.

Anatomical Labeling of DTI Effects

In order to characterize the spatial distribution of effects, we created binary masks of white matter regions representing the corpus callosum (genu, body, splenium), internal capsule (anterior limb, posterior limb), external capsule and anterior, superior and posterior aspects of corona radiata based on the JHU ICBM-DTI-81 WM labels atlas [Mori et al., 2008] provided with FSL. Further, binary masks of select white matter tracts were generated using a probabilistic tractography atlas (JHU WM tractography atlas) [Hua et al., 2008; Mori et al., 2005; Wakana et al., 2004] with a probability threshold of 5% chosen to accommodate inter-subject variation in gross fiber architecture. Eight major WM tracts in each hemisphere (anterior thalamic radiation, cingulum-cingulate gyrus, cingulum-hippocampal gyrus, cortico-spinal tract, inferior and superior longitudinal fasciculus, inferior occipito-frontal fasciculus, and uncinate fasciculus) and two commissural tracts (forceps minor and forceps major) were chosen. Intersections between the WM tract skeleton and the anatomical masks (WM regions and tracts) were used for anatomical labeling of effects from TBSS.

RESULTS

Longitudinal Changes Across Groups

Figure 3 shows the results from the whole-brain voxel-wise analysis of MD changes across groups. Significant ($P < 0.05$, corrected) longitudinal increases in MD were seen in 27.7% of the voxels in the WM skeleton. The effects were distributed in four clusters. The largest cluster exhibited a bilateral frontal distribution (peak-voxel in left anterior external capsule, in proximity to the left uncinate fasiculus). In the right hemisphere, the cluster extended posteriorly along the corona radiata projecting to right parietal and occipital lobules. The three smaller clusters were located in posterior regions of the brain. The largest of these was located in proximity to the precuneus with

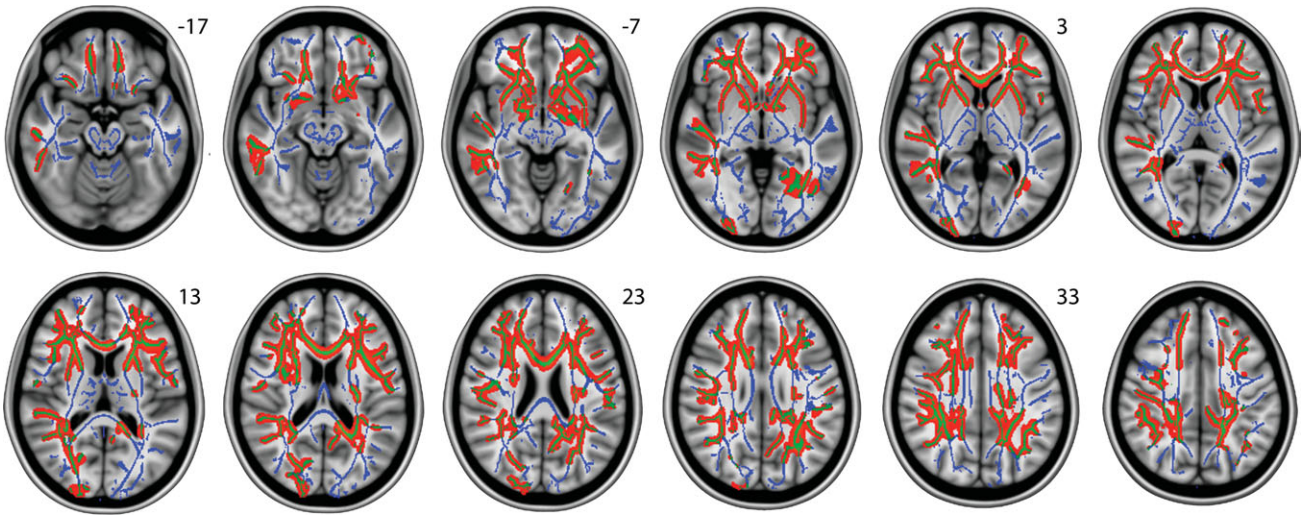

Figure 3.

Longitudinal increase in MD (tp2 – tp1) across groups ($n = 41$). Significant effects ($P < 0.05$, corrected for multiple comparisons) are inflated and displayed in red-green as they appear in the WM skeleton shown in blue. The horizontal background images are corresponding sections of the T1-weighted MNI template brain shown in radiological convention. Numbers denote the Montreal Neurological Institute Z-coordinates.

the peak-voxel overlapping with posterior portions of the probabilistic left anterior thalamic radiation. Additional clusters were near the right occipital pole and the left lateral occipital lobe, respectively.

A quantitative evaluation of the statistical *t*-map suggested that the effects displayed in Figure 3 followed a gradient along the *y*-axis with the strongest effects located in frontal regions of the WM-skeleton (see Fig. 4).

Because of the wide extent of the clusters in Figure 3, we applied a more conservative *P* threshold ($P < 0.025$, corrected) to detect more localized effects. This resulted in one frontal cluster in the left hemisphere (cluster size: 4,284 voxels, MNI-coordinates of peak-voxel [$X = -26$, $Y = 19$, $Z = 3$], located in proximity to the left uncinate fasciculus, see Fig. 5).

Mean MD change in this frontal cluster was 1.3×10^{-5} $mm^2\ s^{-1}$ (SD = 1.5×10^{-5} $mm^2\ s^{-1}$) or 1.6 (2.0)% increase across groups. The rate of change in MD showed a positive correlation with age; suggesting larger MD increases with advancing age (Spearman's $\rho = 0.382$, $P = 0.014$). No significant decreases in FA across groups, or within subgroups, were found in the whole-brain analysis.

Memory Performance

Table II shows group means and standard deviations for the recognition and source memory scores at each timepoint. Repeated measures ANOVA revealed a significant group × time interaction (Greenhouse-Geisser corrected, *F*

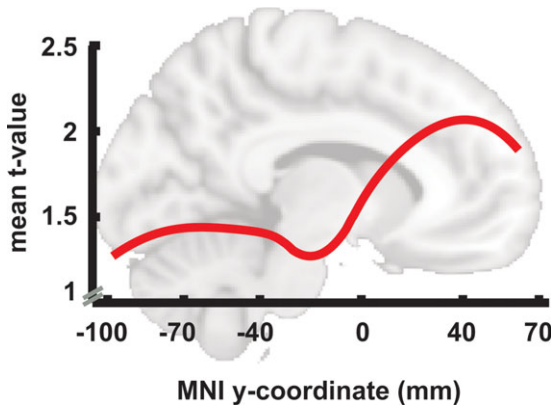

Figure 4.

Anterior–posterior effect gradient. The figure shows an rLOESS curve fit (red) for mean *t*-values at each integer MNI y-coordinate. The distribution of mean *t*-values after testing for longitudinal changes in MD suggests an increasing gradient from posterior to anterior segments. All WM-skeleton *t*-values were first extracted from the raw statistical *t*-map output testing for MD increases over time across all participants. *T*-values at each *y*-coordinate were averaged and analyzed using rLOESS curve fitting in Matlab. A sagittal section of the MNI-template T1-brain is used as background and aligned so that the anatomy corresponds to the respective MNI *y*-coordinates on the *x*-axis of the diagram. Note that neither the *y*-axis nor the *x*-axis of the diagram starts at the conventional 0 levels.

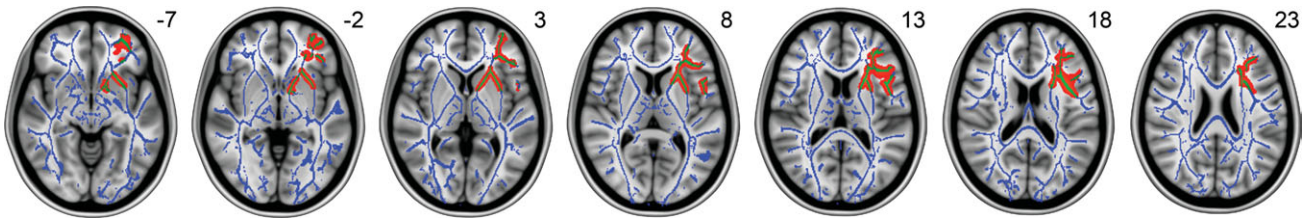

Figure 5.

Cluster used for group × time interaction analysis. After applying a more conservative *P*-threshold to the results indicating longitudinal increase in MD across groups, a cluster in the left anterior hemisphere remained. Significant effects ($P < 0.025$, corrected) are inflated and displayed in redgreen as they appear in the WM skeleton shown in blue. The horizontal background images are corresponding sections of the T1-weighted MNI template brain shown in radiological convention. Numbers denote the Montreal Neurological Institute *Z*-coordinates.

(1, 39) = 14.9, $P < 0.001$). Follow-up longitudinal analyses for each group separately 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, $t = 6.90$, df = 20, $P < 0.0001$), while a small improvement (mean difference = 0.05, SD = 0.12) in the control group 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 ($t = 3.87$, $P < 0.001$). 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 ($P = 0.008$).

Impact of Memory Training on Frontal WM

We hypothesized that memory training would impact the rate of change in DTI-indices of frontal WM. Specifically, we tested for group × time interactions on FA and MD in the frontal cluster showing increase in MD (see Fig. 5). However, we first performed a voxel-wise analysis testing for group differences in MD and FA at tp1. This analysis revealed regions with higher baseline MD in the training group in the left hemisphere, partly overlapping the frontal cluster showing effects of time across groups (Supporting Information Fig. 1, corrected $P < 0.05$), but no differences in FA.

To exclude that baseline differences between groups influenced longitudinal group effects, we included baseline FA and MD skeleton maps as additional voxel-wise regressors in the interaction analysis. When controlling for baseline values, no significant group × time interactions on MD were found. However, the analysis yielded increased FA in the training group compared with controls (corrected $P < 0.05$, cluster size = 410 voxels, Fig. 6a). MNI coordinates at peak voxel were [$X = -21$, $Y = 15$, $Z = 12$], corresponding to the left anterior thalamic radiation (74% probability). Mean FA in this voxel is plotted for each group per time point in Figure 6b. Paired samples *t* tests indicated that the effect was driven by a significant

TABLE II. Group means and standard deviations for memory performance at tp1 and tp2

	Baseline (tp1)		Retest (tp2)		
	M	SD	M	SD	*P*-value*
Memory training (11F/10M)					
Recognition, correct hits	54.0	4.08	53.7	3.89	0.832
Recognition, false alarms	3.24	5.34	1.33	3.14	0.008
Correct source memory judgements	28.0	9.22	39.2	10.0	<0.001
Source/recognition, ratio	0.515	0.149	0.728	0.164	<0.001
Controls (10F/9M)					
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 judgments	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.

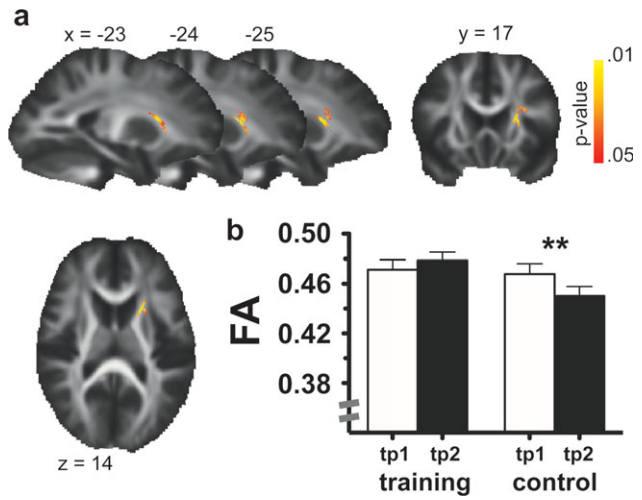

Figure 6.

Impact of memory training on left anterior WM. (**a**) Changes in FA in the left anterior hemisphere in response to cognitive intervention. The effects were estimated by comparing the training and control groups by means of voxel-wise general linear models (group × time interaction). The results are shown on the mean FA image of all participants. The sections are displayed in radiological convention and numbers denote the Montreal Neurological Institute XYZ-coordinates. Red-yellow areas indicate voxels with significant ($P < 0.05$, corrected) relative FA increases in trainers compared with controls. (**b**) Bar plots show mean FA (± 1 SEM) for the peak voxel at scan 1 and 2 for the control and training group. Asterisk (*) indicates $P < 0.05$ based on two-tailed paired samples t tests comparing FA at baseline and at the end of the study phase.

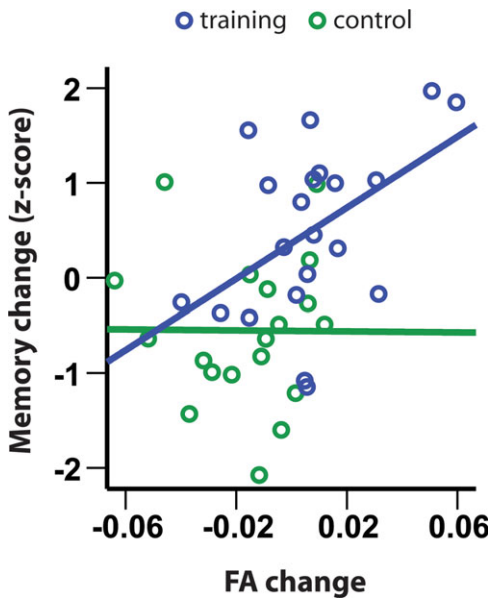

Figure 7.

FA change predicts memory change in the training, but not in the control group, respectively. Scatter plot showing verbal source memory change (Z-scores) as a function of FA change at the peak voxel in Figure 6a. Blue and green lines represent linear fit-lines for the training and control groups, respectively.

decrease in FA in the control group ($t = -3.6$, $P = 0.0019$, df = 19). A trend FA increase was found in the training group, but the result failed to reach statistical significance ($t = 1.4$, $P = 0.17$, df = 20). An independent samples t test on mean FA within the cluster yielded no significant group differences at tp1 ($t = 1.3$, $P = 0.21$, df = 39). No correlations between age and change in FA were found.

The 410 voxels in the cluster showing a significant group × time interaction on FA encompassed four probabilistic association tracts: Left anterior thalamic radiation (216 voxels, 52.7% of total voxels in cluster), inferior fronto-occipito fasciculus (195 voxels, 47.6%), uncinate fasciculus (164 voxels, 40.0%), and superior longitudinal fasciculus (30 voxels, 7.3%), respectively. Note that overlaps exist between the various probabilistic tracts, thus resulting in a sum of more than 100%.

We further examined the cluster showing a significant group × time interaction effect on FA by follow-up analyses on RD and AD in the peak voxel within the cluster. Mean RD for the training and control group was 0.533 10^{-3} mm^2 s^{-1} (standard error (SE) = 0.011 × 10^{-3} mm^2 s^{-1}) and 0.514 (SE = 0.010), respectively, at tp1, and 0.533 (SE = 0.010) and 0.536 (SE = 0.008), respectively, at tp2. Repeated measures ANOVA on RD and group, testing whether mean change in RD differed between groups, revealed a significant group × time interaction (Greenhouse-Geisser corrected, F (1, 39) = 7.11, $P = 0.011$), suggesting larger RD increase in the control group. The same analysis revealed a significant main effect of time (Greenhouse-Geisser corrected, F (1, 39) = 6.86, $P = 0.012$), but no significant main effect of group (Greenhouse-Geisser corrected, F (1, 39) = 0.33, $P = 0.57$). Paired samples t tests revealed a significant increase in RD in controls ($t = 3.90$,

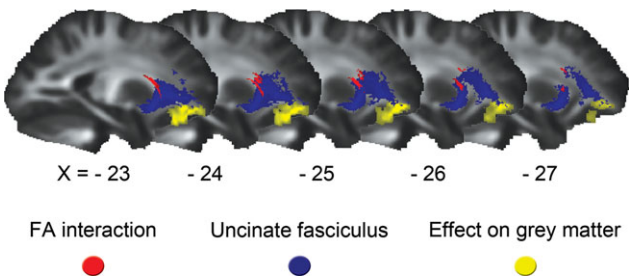

Figure 8.

The anatomical relationship between (1) the probabilistic left uncinate fasciculus, (2) the group × time interaction effects on FA, and (3) the effect in cortical grey matter reported previously [Engvig et al., 2010]. The effects are overlaid left hemisphere sagittal sections of the mean FA image made from all participants.

$P = 0.001$, df = 19), whereas no change was found in the training group ($t = 0.31$, $P = 0.98$, df = 20).

Mean AD for the training and control group was 1.176 $\times 10^{-3}$ mm^2 s^{-1} (SE = 0.017 $10^{-3} \times$ mm^2 s^{-1}), and 1.134 (SE = 0.012), respectively, at tp1, and 1.192 (SE = 0.019) and 1.140 (SE = 0.013), respectively, at tp2. Repeated measures ANOVA on AD and group revealed a significant main effect of time (Greenhouse-Geisser corrected, F (1,39) = 4.22, $P < 0.047$), but no group × time interaction (Greenhouse-Geisser corrected, F (1,39) = 0.84, $P = 0.37$), indicating that AD increased from tp1 to tp2 across groups.

Associations Between Memory and DTI Changes

We assessed relationships between longitudinal DTI changes and memory performance changes by extracting DTI-values from the peak voxel in the cluster showing a group × time interaction effect on FA.

In the training group, FA change correlated positively with change in the memory score targeted by the intervention (correct source judgments at tp2 - correct source judgments at tp1/recognition at tp2 - recognition at tp1) (Spearman's $\rho = 0.496$, $P = 0.022$, $n = 21$). This indicates larger FA increases with increasing memory improvement. Follow-up analysis revealed a significant negative relationship with RD (Spearman's $\rho = -0.547$, $P = 0.010$), indicating decreasing RD with increasing memory improvement. No relationship was found for AD (Spearman's $\rho = -0.026$, $P = 0.910$).

No significant correlations were found in the control group (all P-values >0.42), suggesting that the association was specific to the training group. To illustrate the relationship between FA change and behavior, we plotted the difference in memory performance for each participant as a function of longitudinal FA change at the peak voxel in the left hemisphere (see Fig. 7).

DISCUSSION

We have demonstrated short-term longitudinal changes in WM microstructure as revealed by DTI in middle-aged and elderly healthy volunteers, and further that these alterations were modulated by an intensive memory-training regimen. There were three novel findings: (1) Significant short-term increases in MD with larger effects in anterior compared to posterior regions, (2) the trajectory of anterior FA was influenced by the memory training regimen, and (3) rate of change in anterior FA was positively correlated with verbal memory improvement, suggesting a relationship between behavioral change and microstructural alterations.

Our findings suggest that DTI can be used to detect WM alterations after a period of less than 3 months, demonstrating sensitivity of DTI as an imaging biomarker for microstructural short-term changes in aging. To the best of our knowledge, this is the first study demonstrating an impact of memory training on WM-microstructure. FA, previously reported to be sensitive to alterations in WM in response to motor, sensory and cognitive learning [Hanggi et al., 2010; Keller and Just, 2009; Lövdén et al., 2010; Scholz et al., 2009; Takeuchi et al., 2010; Taubert et al., 2010], showed a significant group × time interaction. The memory-training group showed less negative changes in FA compared with controls. The positive relationship between FA change and change in memory performance supports the neurocognitive significance of the observed effects. The link between brain and behavioral change is important, as cognitive intervention has been proposed as a useful remediation tool in clinical settings [Buschert et al., 2010].

Short-Term Alterations in WM Microstructure in Aging

Our findings indicate that alterations in aging WM, as reflected by the mean diffusivity metric, can be detected not only over years, but also in a matter of months. The direction and extent of changes are in line with the longitudinal MD increases observed in older adults over 2 years in a recent trial [Charlton et al., 2010], and with cross-sectional estimates [Burzynska et al., 2010; Westlye et al., 2010]. The spatial distribution of effects (see Fig. 4) is consistent with an anterior-posterior gradient of WM degradation in aging [Bennett et al., 2010; Davis et al., 2009; Head et al., 2004; Salat et al., 2005]; [but see Barrick et al., 2010; Giorgio et al., 2010]. Further, advancing age was modestly associated with larger MD increases, supporting previous cross-sectional reports of accelerating changes in diffusivity with advancing age [Hasan et al., 2009; Kennedy and Raz, 2009b; Westlye et al., 2010].

We found no significant decreases in FA in the whole brain analysis across groups, in line with a recent 3-month longitudinal study of 25 healthy controls [Mielke et al., 2009]. An alternative interpretation of the lack of FA decrease in the whole-brain analysis in the current study is that the intervention influenced the FA trajectory in the training group, thus in part reducing the overall slope of FA decrease across groups. However, a separate whole-brain analysis on FA decrease in the control group only revealed no significant effects.

This suggests that MD might be more sensitive to short-term changes in aging. In a recent study [Zhang et al., 2010, see Fig. 2 on page 1994], the authors reported widespread regions of age-related MD increase, without significant effects on FA, and pointed out that one interpretation of this is that degenerative brain changes in aging other than demyelination increases the overall water diffusivity while maintaining the underlying directional structure. The reported diffusivity increases give new insight into the short-term dynamics of aging WM in vivo, but the mechanisms underlying these findings are unknown. Post-mortem data suggests that cerebral WM is the structure of

the brain that declines most with age [Marner et al., 2003], with association areas being particularly vulnerable [Bartzokis, 2004]. In addition to demyelination and oligodendrocyte breakdown, expansion of extracellular space [Meier-Ruge et al., 1992] and axon degradation [Marner et al., 2003] accompany normal aging. Future studies in animal models of aging, in which imaging and histological measures can be taken in the same individuals, will be useful to probe the biological basis of the reported changes.

The Impact of Memory Training on Longitudinal Changes

We investigated the impact of memory training on short-term changes in MD and FA. A group × time interaction analysis revealed frontal areas of FA reduction in the control group where a training group showed no changes. This might suggest a positive impact of the intervention in line with previous reports. It was recently shown that about half an hour working memory training a day for two months was associated with increased FA in frontal regions in young adults [Takeuchi et al., 2010]. Moreover, up to 5% increase in FA after a 6-week motor training intervention has been documented [Scholz et al. 2009].

In the present study, follow-up analyses revealed that the memory training group showed less RD increase in frontal WM compared with controls, whereas changes in AD were of similar magnitude in both groups. These findings are compatible with the pattern reported in Lövdén et al. [2010], suggesting intact WM plasticity in response to cognitive training in adulthood.

Selective changes in RD have partly been linked to myelination [Song et al., 2005], although the neurobiological underpinnings of changes in diffusion parameters are still not understood [see e.g. Concha et al., 2010]. The concurrent increases in AD in the present study also complicate the interpretation of underlying neuronal mechanisms [Wheeler-Kingshott and Cercignani, 2009]. Regardless of these limitations, recent research suggest a role of myelin in cognitive function and learning, and that the myelination process is modifiable by experience [for a review, see Fields, 2008]. Experiments in cell cultures indicate that neuronal activity promotes myelination [Ishibashi et al., 2006], and blocking neuronal firing inhibits myelination over short time-spans [Demerens et al., 1996]. A recent study of spine formation after skill-learning in rats further indicates that the mechanisms underlying neuronal adaptation to learning are rapid phenomena, happening within days [Xu et al., 2009]. Thus, it is not unlikely that cognitive training can have effects on myelination and myelination-related processes, and that the observed changes in diffusion parameters are related to these. However, any interpretation of diffusion changes must be made with caution at this stage due to our incomplete understanding of the relationship between DTI and WM microstructure [Beaulieu, 2009].

The effect of training on FA changes in the present study was mainly driven by a significant decrease in the control group, which is consistent with cross-sectional DTI-investigations of normal aging [Bennett et al., 2010; Head et al., 2004; Westlye et al., 2010]. One recent longitudinal study with a 1-year scan interval reported FA decreases of up to 10%, including regions overlapping the inferior fronto-occipito fasciculus and the uncinate in healthy controls [Teipel et al., 2010].

It should be noted that the magnitude of FA decrease (Fig. 6a) in the control group was relatively large (mean decrease of 3.6% at peak voxel). Recruiting participants to memory training could possibly attract volunteers seeking ways to prevent or halt subtle symptoms of cognitive decline and neurodegeneration. However, as reported in the methods section, this was thoroughly screened for, and due to the randomization procedure it should not be different among groups.

To exert maximum control over confounding factors we employed a randomized controlled design. There were no baseline differences between groups in any demographic variables or neuropsychological measures. Despite this, we found regional baseline group differences in MD, with higher MD-values in the training group, somewhat overlapping the frontal area showing effect of time. No significant baseline differences were found for FA. This implies that the three eigenvalues were proportionately higher in the training group in these regions—an occurrence that, per definition, will not be reflected by the FA metric.

We controlled for this baseline difference when assessing training-related effects by using tp1-values as voxelwise regressors. Statistically controlling for baseline level resulted in no significant group × time interaction for MD. It is noteworthy that the absolute difference between groups in MD (0.021×10^{-3} mm^2 s^{-1}) at tp1 in the frontal cluster (see Fig. 5) is comparable to that of a previous study [Lövdén et al., 2010] (MD group difference in corpus callosum at baseline, segment 1: 0.031×10^{-3} mm^2 s^{-1}). Unfortunately, this issue was not specifically addressed in that study, and so it remains uncertain what impact statistically controlling for such baseline differences might have. In the present study, a significant effect of training on FA changes was still found when tp1 values were accounted for. However, even though controlled for by a conventional statistical approach, it is difficult to be certain that baseline differences did not impact the results. Caution is warranted when interpreting neuroimaging studies of brain plasticity, and we argue that large randomized controlled trials are needed in order to better clarify the impact of cognitive training on aging brain structure.

In the present study FA and MD yielded complementary results. MD was more sensitive to both longitudinal across-group changes, and between group differences at baseline. On the other hand, FA-changes within regions showing longitudinal increases in overall diffusivity

differed between groups. Furthermore, FA-changes correlated with a measure of behavioral change. These observations exemplify how two indices of the diffusion tensor yield complementary results that are in turn likely to reflect different biological phenomena. FA is often regarded as a measure of intravoxel coherence of the white matter microstructure [e.g., Pfefferbaum et al., 2000]. Any biological structure hindering free water diffusion disproportionately with regards to the axes of the tensor model, can impact the degree of FA [Beaulieu, 2009]. MD does not consider orientation and instead serves as an overall measure for barrier sparseness.

Associations Between Memory and DTI Changes

We found a relationship between changes in FA and changes in memory performance as specifically targeted by the MoL. This indicates more positive alterations in subjects benefiting the most from training. Studies of the relationship between FA and behavior have reported both negative [Hanggi et al., 2010; Imfeld et al., 2009] and positive associations [Bengtsson et al., 2005] with behavioral practice in young adults. However, the positive relationship between FA and a cognitive outcome reported here is in line with DTI and cognition studies in middle and late parts of the life span [Kennedy and Raz, 2009a; O'Sullivan et al., 2001a].

The association was specific for the intervention group, which is in line with the notion of specific cognitive training related effects. Interestingly, the cluster showing an effect of memory training on FA encompassed important association fibers, particularly the anterior thalamic radiation, and the inferior occipito-frontal and uncinate fasciculi in the left hemisphere. The fact that several association fibers overlapped with this cluster, suggest that the effect was found in a relatively busy WM region including crossing fibers. FA values were also relatively low as compared to more coherent WM regions (see Fig. 6b).

The anterior thalamic radiation projects from nuclei in anterior and medial regions of the thalamus suggested to be involved in declarative memory function [Van der Werf et al., 2003]. The inferior occipito-frontal and uncinate fasciculi have putative roles in semantic processing [Duffau et al., 2005] and memory [Papagno et al., 2011], respectively. Verbal memory formation is further shown to involve the left prefrontal hemisphere [Elmer et al., 2009; Wagner et al., 1998]. This supports the observed training-related changes, since the present intervention aimed at training and improving verbal memory function through learning and serial recall of word-lists.

The brain-behavior relationships suggest that both changes in fractional anisotropy and also radial diffusivity in the left frontal region predict improvement in verbal source memory, targeted by the intervention. Given that the observed changes in fiber tract integrity are coupled to specific verbal memory improvement, contributions of encoding, rehearsal and recollection-related processes involved with practice are likely [Oberauer and Meyer, 2009]. More sophisticated behavioral paradigms could aid future dissection of the active cognitive elements involved in intervention-related brain and behavior changes. Future studies could further try to link these elements with the structural changes by investigating task-specific neuronal activation patterns before and after training by e.g., functional MRI or quantitative EEG.

Nearly 40% of the voxels showing training-related changes in FA encompassed the left uncinate fasciculus. A recent study demonstrated associations between uncinate FA and several indices of episodic and working memory performance [Kennedy and Raz, 2009a]. Moreover, Flöel et al. [2009] used probabilistic tractography to show that FA in WM tracts arising from Broca's area projecting to the anterior temporal lobe via the uncinate predicted grammar learning success. This indicates relevance of FA with regards to functional specificity of these tracts. Interestingly, the uncinate projects to the orbitofrontal cortex [Schmahmann et al., 2007], for which we have recently reported training-related changes in GM after the present intervention in the same sample [Engvig et al., 2010]. Figure 8 illustrates the anatomical relationships between the probabilistic left uncinate fasciculus, the group $\times$ time interaction effect for FA, and the GM effect reported previously. Convergent findings across imaging modalities provide a valuable validation of the results, and might point to a common neurobiological mechanism of cognitive training affecting both cerebral GM and WM.

Summarized, our findings suggest that DTI indices of WM microstructure can be directly related to behavioral measures of cognition and thus point to the usefulness of longitudinal DTI assessments in relation to clinical outcomes. This is supported by accumulating evidence indicating that individual differences in DTI indices are functionally [Saur et al., 2008; Westlye et al., 2009] and behaviorally [Bengtsson et al., 2005; Della-Maggiore et al., 2009; Johansen-Berg et al., 2007; Oechslin et al., 2009] relevant [for reviews, see Chanraud et al., 2010; Johansen-Berg, 2010]. However, despite the recent progress in linking imaging and cognitive measures, further studies delineating the complex relationships between brain function, structure and behavior are highly needed.

Methodological Considerations

The present study has several methodological strengths. First, data were acquired using 30 diffusion directions providing higher angular resolution and thus more robust estimations of the diffusion tensor than several previous DTI studies of WM plasticity. Second, to account for residual variation due to cross-subject comparisons we used TBSS, an automated technique for sensitive, robust and unbiased voxel-wise analysis of DTI data [Smith et al., 2006], in combination with a highly sensitive approach to

statistical thresholding [Smith and Nichols, 2009]. Third, to account for residual spatial variation due to multiple time-points we used a robust longitudinal analysis protocol employing validated spatial alignment procedures designed for longitudinal analysis [Smith et al., 2002]. In the present study our approach assured excellent anatomical alignment across time-points, further validating our findings.

We employed a randomized controlled design, with age-, sex-, and education-matched controls, which should enable dissociation of intervention-related changes from other factors affecting the age-related trajectories. Further, we used strict exclusion criteria to avoid the inclusion of subjects with conditions known to interfere with WM microstructure. It is possible that recruiting older adults to memory training through newspaper ads could attract subjects with preclinical cognitive impairment searching for ways to alleviate their memory concerns. Thus, we carefully excluded individuals with any history of brain diseases, neuropsychological deviations or abnormalities on MRI as evaluated by an experienced neuroradiologist. We paid particular attention to severe hypertension, diabetes mellitus, cardiovascular diseases and stroke, since these conditions are relatively common among older individuals and can result in degradation of WM microstructure [O'Sullivan et al., 2001b]. Furthermore, all participants specifically reported not to be concerned about their own memory function. However, we did not differentiate between normal appearing white matter and mild white matter lesions in our DTI-analyses, although we screened the participants for absence of major cerebrovascular disease and major lesions. Thus a limitation of the present study is that in addition to other age-related processes, mild white matter lesions might underlie some of the regional findings of DTI-alterations. However, as the participants were randomly assigned to the training or control group, we do not believe that this may have impacted the results of the cognitive intervention.

In common with most studies utilizing convenience samples, the age- and sex-adjusted IQ and CVLT scores indicated that the current sample included participants with relatively high cognitive abilities (see Methods). Our results thus may not generalize to a random community sample with lower education and poorer cognitive performance. Other limitations to the present study include the use of short-term assessments only. Further studies are needed in order to establish the long-term cognitive and cerebral effects of cognitive training. More observations are needed to assess or model the temporal aspects of the training-related impact on WM trajectories. As discussed above, the biological underpinnings of changes in DTI indices are unclear, even though some efforts have been made to validate DTI with histology in patients [Concha et al., 2010]. Future studies should incorporate DTI and histology, for example by training rodents and examining the short-term changes using both MRI and invasive techniques.

Further, the use of a randomized, controlled design should ideally include an active control condition to better account for the psychosocial components of a group-based training program [see Fratiglioni et al., 2004, for a review on psychosocial effects on cognition in late life]. The fact that we found baseline differences in MD support the need of larger samples in order to properly delineate effects of cognitive training in aging. In the present study we opted to statistically control for baseline values. Regression toward the mean is a common phenomenon in longitudinal studies [Barnett et al., 2005] which complicates the interpretations of experience-dependent plasticity [Mozolic et al., 2010].

CONCLUSION

In conclusion, our results demonstrate that longitudinal changes in WM microstructure can be detected over an interval of only 10 weeks in middle-aged and older adults. Further, WM changes differed between the intervention and control groups, suggesting a positive impact of memory training on short-term WM-trajectories. Change in memory performance specifically addressed by the intervention correlated positively with the rate of change in FA, adding to the accumulating evidence supporting DTI as a behaviorally relevant imaging biomarker. Although encouraging, great caution is warranted when making biological interpretations from imaging biomarkers, and larger randomized controlled studies of the effects of cognitive training interventions on aging brain structure are highly needed.

ACKNOWLEDGMENTS

The authors thank Paulina Due-Tønnessen, MD, for performing the neuroradiological evaluations. The authors declare no competing financial interests.

REFERENCES

Barnett AG, van der Pols JC, Dobson AJ (2005): Regression to the mean: What it is and how to deal with it. Int J Epidemiol 34:215–220.

Barrick TR, Charlton RA, Clark CA, Markus HS (2010): White matter structural decline in normal ageing: A prospective longitudinal study using tract-based spatial statistics. NeuroImage 51:565–577.

Bartzokis G (2004): Age-related myelin breakdown: A developmental model of cognitive decline and Alzheimer's disease. Neurobiol Aging 25:5–18; author reply 49–62.

Beaulieu C (2009): The biological basis of diffusion anisotropy. In: Johansen-Berg H, Behrens TEJ, editors. Diffusion MRI: From Quantitative Measurement to In-Vivo Neuroanatomy. London: Elsevier. pp 105–126.

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

Bennett IJ, Madden DJ, Vaidya CJ, Howard DV, Howard JH Jr. (2010): Age-related differences in multiple measures of white matter integrity: A diffusion tensor imaging study of healthy aging. Hum Brain Mapp 31:378–390.

Bower GH (1970): Analysis of a mnemonic device. Am Scientist 58:496–510.

Burzynska AZ, Preuschhof C, Bäckman L, Nyberg L, Li SC, Lindenberger U, Heekeren HR (2010): Age-related differences in white matter microstructure: Region-specific patterns of diffusivity. NeuroImage 49:2104–2112.

Buschert V, Bokde ALW, Hampel H (2010): Cognitive intervention in Alzheimer disease. Nat Rev Neurol 6:508–517.

Chanraud S, Zahr N, Sullivan E, Pfefferbaum A (2010): MR diffusion tensor imaging: A window into white matter integrity of the working brain. Neuropsychol Rev 20:209–225.

Charlton RA, Schiavone F, Barrick TR, Morris RG, Markus HS (2010): Diffusion tensor imaging detects age related white matter change over a 2 year follow-up which is associated with working memory decline. J Neurol Neurosurg Psychiatry 81:13–19.

Cleveland WS, Devlin SJ (1988): Locally weighted regression: An approach to regression analysis by local fitting. J Am Stat Assoc 83:596–610.

Concha L, Livy DJ, Beaulieu C, Wheatley BM, Gross DW (2010): In vivo diffusion tensor imaging and histopathology of the fimbria-fornix in temporal lobe epilepsy. J Neurosci 30:996–1002.

Davis SW, Dennis NA, Buchler NG, White LE, Madden DJ, Cabeza R (2009): Assessing the effects of age on long white matter tracts using diffusion tensor tractography. NeuroImage 46:530–541.

Delis DC, Kramer JH, Kaplan E, Ober BA. (2000): California Verbal Learning Test, 2nd ed (CVLT-II). San Antonio, TX: The Psychological Corporation.

Della-Maggiore V, Scholz J, Johansen-Berg H, Paus T (2009): The rate of visuomotor adaptation correlates with cerebellar white-matter microstructure. Hum Brain Mapp 30:4048–4053.

Demerens C, Stankoff B, Logak M, Anglade P, Allinquant B, Couraud F, Zalc B, Lubetzki C (1996): Induction of myelination in the central nervous system by electrical activity. Proc Natl Acad Sci USA 93:9887–9892.

Duffau H, Gatignol P, Mandonnet E, Peruzzi P, Tzourio-Mazoyer N, Capelle L (2005): New insights into the anatomo-functional connectivity of the semantic system: A study using cortico-subcortical electrostimulations. Brain 128:797–810.

Elmer S, Burkard M, Renz B, Meyer M, Jancke L (2009): Direct current induced short-term modulation of the left dorsolateral prefrontal cortex while learning auditory presented nouns. Behav Brain Funct 5:29.

Engvig A, Fjell AM, Westlye LT, Moberget T, Sundseth Ø, Larsen VA, Walhovd KB (2010): Effects of memory training on cortical thickness in the elderly. NeuroImage 52:1667–1676.

Fields RD (2008): White matter in learning, cognition and psychiatric disorders. Trends Neurosci 31:361–370.

Fjell AM, Walhovd KB, Fennema-Notestine C, McEvoy LK, Hagler DJ, Holland D, Brewer JB, Dale AM (2009): One-year brain atrophy evident in healthy aging. J Neurosci 29:15223–15231.

Flöel A, de Vries MH, Scholz J, Breitenstein C, Johansen-Berg H (2009): White matter integrity in the vicinity of Broca's area predicts grammar learning success. NeuroImage 47:1974–1981.

Folstein MF, Folstein SE, McHugh PR (1975): "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. J Psychiatric Res 12:189–198.

Fratiglioni L, Paillard-Borg S, Winblad B (2004): An active and socially integrated lifestyle in late life might protect against dementia. Lancet Neurol 3:343–353.

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

Hanggi J, Koeneke S, Bezzola L, Jancke L (2010): Structural neuroplasticity in the sensorimotor network of professional female ballet dancers. Hum Brain Mapp 31:1196–1206.

Hasan KM, Kamali A, Iftikhar A, Kramer LA, Papanicolaou AC, Fletcher JM, Ewing-Cobbs L (2009): Diffusion tensor tractography quantification of the human corpus callosum fiber pathways across the lifespan. Brain Res 1249:91–100.

Head D, Buckner RL, Shimony JS, Williams LE, Akbudak E, Conturo TE, McAvoy M, Morris JC, Snyder AZ (2004): Differential vulnerability of anterior white matter in nondemented aging with minimal acceleration in dementia of the Alzheimer type: Evidence from diffusion tensor imaging. Cerebral Cortex 14:410–423.

Hua K, Zhang J, Wakana S, Jiang H, Li X, Reich D, Calabresi P, Pekar J, van Zijl P, Mori S (2008): Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification. Neuroimage 39:336–347.

Imfeld A, Oechslin MS, Meyer M, Loenneker T, Jancke L (2009): White matter plasticity in the corticospinal tract of musicians: A diffusion tensor imaging study. NeuroImage 46:600–607.

Ishibashi T, Dakin KA, Stevens B, Lee PR, Kozlov SV, Stewart CL, Fields RD (2006): Astrocytes promote myelination in response to electrical impulses. Neuron 49:823–832.

Jenkinson M, Smith S (2001): A global optimization method for robust affine registration of brain images. Med Image Analysis 5:143–156.

Jennings JM, Webster LM, Kleykamp BA, Dagenbach D (2005): Recollection training and transfer effects in older adults: Successful use of a repetition-lag procedure. Aging Neuropsychol Cogn 12:278–298.

Johansen-Berg H (2010): Behavioural relevance of variation in white matter microstructure. Curr Opin Neurol 23:351–358.

Johansen-Berg H, Della-Maggiore V, Behrens TEJ, Smith SM, Paus T (2007): Integrity of white matter in the corpus callosum correlates with bimanual co-ordination skills. NeuroImage 36 (Supp 2):T16–T21.

Jäncke L (2009): The plastic human brain. Restor Neurol Neurosci 27:521–538.

Keller TA, Just MA (2009): Altering cortical connectivity: Remediation-induced changes in the white matter of poor readers. Neuron 64:624–631.

Kennedy KM, Raz N. (2009a):Aging white matter and cognition: Differential effects of regional variations in diffusion properties on memory, executive functions, and speed. Neuropsychologia 47:916–927.

Kennedy KM, Raz N. (2009b):Pattern of normal age-related regional differences in white matter microstructure is modified by vascular risk. Brain Res 1297:41–56.

Kliegl R, Smith J, Baltes PB (1990): On the locus and process of magnification of age differences during mnemonic training. Dev Psychol 26:894–904.

Le Bihan D (2003): Looking into the functional architecture of the brain with diffusion MRI. Nat Neurosci 4:469–480.

Lustig C, Shah P, Seidler R, Reuter-Lorenz P (2009): Aging, training, and the brain: A review and future directions. Neuropsychol Rev 19:504–522.

Lövdén M, Bodammer NC, Kühn S, Kaufmann J, Schütze H, Tempelmann C, Heinze H-J, Düzel E, Schmiedek F, Lindenberger U (2010): Experience-dependent plasticity of white-matter microstructure extends into old age. Neuropsychologia 48:3878–3883.

Madden DJ, Whiting WL, Huettel SA, White LE, MacFall JR, Provenzale JM (2004): Diffusion tensor imaging of adult age differences in cerebral white matter: Relation to response time. Neuroimage 21:1174–1181.

Marner L, Nyengaard JR, Tang Y, Pakkenberg B (2003): Marked loss of myelinated nerve fibers in the human brain with age. J Comp Neurol 462:144–152.

Meier-Ruge W, Ulrich J, BrüHlmann M, Meier E (1992): Age-related white matter atrophy in the human brain. Ann N Y Acad Sci 673:260–269.

Mielke MM, Kozauer NA, Chan KCG, George M, Toroney J, Zerrate M, Bandeen-Roche K, Wang MC, vanZijl P, Pekar JJ, Mori S, Lyketsos CG, Albert M (2009): Regionally-specific diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease. NeuroImage 46:47–55.

Minati L, Grisoli M, Bruzzone MG (2007): MR spectroscopy, functional MRI, and diffusion-tensor imaging in the aging brain: A conceptual review. J Geriatr Psychiatry Neurol 20:3–21.

Mori S, Wakana S, van Zijl PCM, Nagae-Poetscher LM (2005): MRI Atlas of Human White Matter. Amsterdam: Elsevier.

Mori S, Oishi K, Jiang H, Jiang L, Li X, Akhter K, Hua K, Faria AV, Mahmood A, Woods R, Toga AW, Pike GB, Neto PR, Evans A, Zhang J, Huang H, Miller MI, van Zijl P, Mazziotta J (2008): Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. NeuroImage 40:570–582.

Mozolic JL, Hayaska S, Laurienti PJ (2010): A cognitive training intervention increases resting cerebral blood flow in healthy older adults. Front Hum Neurosci 4:12.

Murphy EA, Holland D, Donohue M, McEvoy LK, Hagler DJ Jr, Dale AM, Brewer JB (2010): Six-month atrophy in MTL structures is associated with subsequent memory decline in elderly controls. NeuroImage 53:1310–1317.

Nichols TE, Holmes AP (2002): Nonparametric permutation tests for functional neuroimaging: A primer with examples. Hum Brain Mapp 15:1–25.

O'Sullivan M, Jones DK, Summers PE, Morris RG, Williams SC, Markus HS (2001a): Evidence for cortical "disconnection" as a mechanism of age-related cognitive decline. Neurology 57:632–638.

O'Sullivan M, Summers PE, Jones DK, Jarosz JM, Williams SCR, Markus HS (2001b): Normal-appearing white matter in ischemic leukoaraiosis: A diffusion tensor MRI study. Neurology 57:2307–2310.

Oberauer K, Meyer N (2009): The contributions of encoding, retention, and recall to the Hebb effect. Memory 17:774–781.

Oechslin MS, Imfeld A, Loenneker T, Meyer M, Jancke L (2009): The plasticity of the superior longitudinal fasciculus as a function of musical expertise: A diffusion tensor imaging study. Front Hum Neurosci 3:76.

Papagno C, Miracapillo C, Casarotti A, Romero Lauro LJ, Castellano A, Falini A, Casaceli G, Fava E, Bello L (2011): What is the role of the uncinate fasciculus? Surgical removal and proper name retrieval. Brain 134:405–414.

Pfefferbaum A, Sullivan EV (2003): Increased brain white matter diffusivity in normal adult aging: Relationship to anisotropy and partial voluming. Magn Reson Med 49:953–961.

Pfefferbaum A, Sullivan EV, Hedehus M, Lim KO, Adalsteinsson E, Moseley M (2000): Age-related decline in brain white matter anisotropy measured with spatially corrected echo-planar diffusion tensor imaging. Magn Reson Med 44:259–268.

Pfefferbaum A, Adalsteinsson E, Sullivan EV (2005): Frontal circuitry degradation marks healthy adult aging: Evidence from diffusion tensor imaging. Neuroimage 26:891–899.

Reese TG, Heid O, Weisskoff RM, Wedeen VJ (2003): Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. Magn Reson Med 49:177–182.

Reichman WE, Fiocco AJ, Rose NS (2010): Exercising the brain to avoid cognitive decline: Examining the evidence. Aging Health 6:565–584.

Salat DH, Tuch DS, Greve DN, van der Kouwe AJW, Hevelone ND, Zaleta AK, Rosen BR, Fischl B, Corkin S, Rosas HD, Dale AM (2005): Age-related alterations in white matter microstructure measured by diffusion tensor imaging. Neurobiol Aging 26:1215–1227.

Saur D, Kreher BW, Schnell S, Kümmerer D, Kellmeyer P, Vry M-S, Umarova R, Musso M, Glauche V, Abel S, Huber W, Rijntjes M, Hennig J, Weiller C (2008): Ventral and dorsal pathways for language. Proc Natl Acad Sci 105:18035–18040.

Schmahmann JD, Pandya DN, Wang R, Dai G, D'Arceuil HE, de Crespigny AJ, Wedeen VJ (2007): Association fibre pathways of the brain: Parallel observations from diffusion spectrum imaging and autoradiography. Brain 130:630–653.

Schmierer K, Wheeler-Kingshott CA, Tozer DJ, Boulby PA, Parkes HG, Yousry TA, Scaravilli F, Barker GJ, Tofts PS, Miller DH (2008): Quantitative magnetic resonance of postmortem multiple sclerosis brain before and after fixation. Magn Reson Med 59:268–277.

Scholz J, Klein MC, Behrens TEJ, Johansen-Berg H (2009): Training induces changes in white-matter architecture. Nat Neurosci 12:1370–1371.

Smith SM (2002): Fast robust automated brain extraction. Hum Brain Mapp 17:143–155.

Smith SM, Nichols TE (2009): Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localization in cluster inference. NeuroImage 44:83–98.

Smith SM, De Stefano N, Jenkinson M, Matthews PM (2001): Normalized accurate measurement of longitudinal brain change. J Comput Assist Tomogr 25:466–475.

Smith SM, Zhang Y, Jenkinson M, Chen J, Matthews PM, Federico A, De Stefano N (2002): Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. NeuroImage 17:479–489.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004): Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23 (Suppl 1):S208–S219.

Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM, Behrens TEJ (2006): Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. NeuroImage 31:1487–1505.

Smith SM, Rao A, De Stefano N, Jenkinson M, Schott JM, Matthews PM, Fox NC (2007): Longitudinal and cross-sectional analysis of atrophy in Alzheimer's disease: Cross-validation of BSI, SIENA, and SIENAX. NeuroImage 36:1200–1206.

Song SK, Sun SW, Ramsbottom MJ, Chang C, Russell J, Cross AH (2002): Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. Neuroimage 17:1429–1436.

Song SK, Yoshino J, Le TQ, Lin SJ, Sun SW, Cross AH, Armstrong RC (2005): Demyelination increases radial diffusivity in corpus callosum of mouse brain. Neuroimage 26:132–140.

Stevens B, Porta S, Haak LL, Gallo V, Fields RD (2002): Adenosine: A neuron-glial transmitter promoting myelination in the CNS in response to action potentials. Neuron 36:855–868.

Sun SW, Liang HF, Schmidt RE, Cross AH, Song SK (2007): Selective vulnerability of cerebral white matter in a murine model of multiple sclerosis detected using diffusion tensor imaging. Neurobiol Dis 28:30–38.

Takeuchi H, Sekiguchi A, Taki Y, Yokoyama S, Yomogida Y, Komuro N, Yamanouchi T, Suzuki S, Kawashima R (2010): Training of working memory impacts structural connectivity. J Neurosci 30:3297–3303.

Taubert M, Draganski B, Anwander A, Muller K, Horstmann A, Villringer A, Ragert P (2010): Dynamic properties of human brain structure: Learning-related changes in cortical areas and associated fiber connections. J Neurosci 30:11670–11677.

Teipel SJ, Meindl T, Wagner M, Stieltjes B, Reuter S, Hauenstein KH, Filippi M, Ernemann U, Reiser MF, Hampel H (2010): Longitudinal changes in fiber tract integrity in healthy aging and mild cognitive impairment: A DTI follow-up study. J Alzheimers Dis 22:507–522.

Van der Werf YD, Jolles J, Witter MP, Uylings HB (2003): Contributions of thalamic nuclei to declarative memory functioning. Cortex 39:1047–1062.

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.

Wagner AD, Schacter DL, Rotte M, Koutstaal W, Maril A, Dale AM, Rosen BR, Buckner RL (1998): Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. Science 281:1188–1191.

Wakana S, Jiang H, Nagae-Poetscher LM, van Zijl PC, Mori S (2004): Fiber tract-based atlas of human white matter anatomy. Radiology 230:77–87.

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

Westlye LT, Walhovd KB, Bjornerud A, Due-Tonnessen P, Fjell AM (2009): Error-related negativity is mediated by fractional anisotropy in the posterior cingulate gyrus—A study combining diffusion tensor imaging and electrophysiology in healthy adults. Cereb Cortex 19:293–304.

Westlye LT, Walhovd KB, Dale AM, Bjornerud A, Due-Tonnessen P, Engvig A, Grydeland H, Tamnes CK, Ostby Y, Fjell AM (2010): Life-span changes of the human brain white matter: Diffusion tensor imaging (DTI) and volumetry. Cereb Cortex 20:2055–2068.

Wheeler-Kingshott CAM, Cercignani M (2009): About "axial" and "radial" diffusivities. Magn Reson Med 61:1255–1260.

Xu T, Yu X, Perlik AJ, Tobin WF, Zweig JA, Tennant K, Jones T, Zuo Y (2009): Rapid formation and selective stabilization of synapses for enduring motor memories. Nature 462:915–919.

Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO (1983): Development and validation of a geriatric depression screening scale: A preliminary report. J Psychiatr Res 17:37–49.

Zhang Y, Du AT, Hayasaka S, Jahng GH, Hlavin J, Zhan W, Weiner MW, Schuff N (2010): Patterns of age-related water diffusion changes in human brain by concordance and discordance analysis. Neurobiol Aging 31:1991–2001.