



## Volumetric and microstructural regional changes of the hippocampus underlying development of recall performance after extended retention intervals

Anders M. Fjell<sup>a,b,\*</sup>, Markus H. Sneve<sup>a</sup>, Donatas Sederevicius<sup>a</sup>, Øystein Sørensen<sup>a</sup>, Stine K. Krogsrud<sup>a</sup>, Athanasia M. Mowinckel<sup>a</sup>, Kristine B. Walhovd<sup>a,b</sup>

<sup>a</sup> Center for Lifespan Changes in Brain and Cognition, University of Oslo, Norway

<sup>b</sup> Department of Radiology and Nuclear Medicine, Oslo University Hospital, Oslo, Norway

### ARTICLE INFO

#### Keywords:

Episodic memory  
Long-term memory  
Hippocampus  
Development  
Volume  
Mean diffusion

### ABSTRACT

Performance on recall tests improves through childhood and adolescence, in part due to structural maturation of the medial temporal cortex. Although partly different processes support successful recall over shorter vs. longer intervals, recall is usually tested after less than an hour. The aim of the present study was to test whether there are unique developmental changes in recall performance using extended retention intervals, and whether these are related to structural maturation of sub-regions of the hippocampus. 650 children and adolescents from 4.1 to 24.8 years were assessed in total 962 times (mean interval  $\approx$  1.8 years). The California Verbal Learning Test (CVLT) and the Rey Complex Figure Test (CFT) were used. Recall was tested 30 min and  $\approx$  10 days after encoding. We found unique developmental effects on recall in the extended retention interval condition independently of 30 min recall performance. For CVLT, major improvements happened between 10 and 15 years. For CFT, improvement was linear and was accounted for by visuo-constructive abilities. The relationships did not show anterior-posterior hippocampal axis differences. In conclusion, performance on recall tests using extended retention intervals shows unique development, likely due to changes in encoding depth or efficacy, or improvements of long-term consolidation processes.

### 1. Introduction

Recall performance improves through childhood and adolescence (Bauer, 2015), likely partly due to structural maturation of critical brain regions such as the medial temporal and the prefrontal cortex (Ostby et al., 2012; DeMaster et al., 2014). Of special importance for the ability to recall previous events is the hippocampus. Hippocampus is necessary for encoding of new information (Scoville and Milner, 1957), and engaged in retrieval of vivid episodic memories (Geib et al., 2017). Hippocampus also plays a unique role in consolidation and maintenance of episodic memories (Moscovitch et al., 2016). Accordingly, studies have shown stronger relationships between hippocampal volume and memory over days and weeks compared to hours or less in development (Ostby et al., 2012), adulthood and aging (Walhovd et al., 2004). Studies of hippocampal-neocortical connectivity also point to stronger involvement of hippocampus for encoding of episodic information that can be successfully recalled after extended retention intervals (Sneve

et al., 2015).

The aim of the present study was to test whether there are unique developmental improvements in recall performance after extended retention intervals, and whether structural development of sub-regions of the hippocampus is related to this. Different hippocampal sub-regions are involved in processing of episodic information, with partly distinguishable contributions (Moscovitch et al., 2016; Strange et al., 2014; Collin et al., 2015). Much focus has been on a long axis anterior-posterior (AP) gradient (Poppenk et al., 2013; Chase et al., 2015; Kühn and Gallinat, 2014), since partly different age-trajectories of structure and activity along the AP-axis have been found (DeMaster et al., 2014; Lin et al., 2013; Riggins et al., 2018; Gogtay et al., 2006; Schlichting et al., 2017; Daugherty et al., 2017). Relationships between recall performance and the volume of anterior-posterior sub-regions have also been reported to differ between children and adults. One study found that the volume of the posterior hippocampus was related to recall performance in children while anterior volume (DeMaster et al., 2014)

\* Corresponding author at: Dept of Psychology, Pb. 1094 Blindern, 0317 Oslo, Norway.

E-mail address: [andersmf@psykologi.uio.no](mailto:andersmf@psykologi.uio.no) (A.M. Fjell).

<https://doi.org/10.1016/j.dcn.2019.100723>

Received 1 April 2019; Received in revised form 10 October 2019; Accepted 19 October 2019

Available online 22 October 2019

1878-9293/© 2019 The Author(s).

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

and activity (Sastre et al., 2016) was related in adults, the latter in accordance with results of a large adult study (Hackert et al., 2002). The literature is not consistent, however, since other studies have found recall performance-volume correlations in the anterior hippocampus in children (Riggins et al., 2015) and the posterior in young adults (Poppenk and Moscovitch, 2011).

Volumetric changes may reflect various cellular processes within the hippocampus, such as neurogenesis (Goncalves et al., 2016), non-neuronal cell changes (Bechmann and Nitsch, 2000), cell death and synaptic changes (Small et al., 2011; Lester et al., 2017), pruning (Kantor and Kolodkin, 2003), myelination (Nickel and Gu, 2018) and vascularization (Tatu and Vuillier, 2014). Several of these processes may alter water diffusion in the tissue, which can be measured by diffusion tensor imaging (DTI). Higher mean diffusivity (MD) in the hippocampus is related to aging (Pereira et al., 2014; den Heijer et al., 2012; Wolf et al., 2015), increases over time in older adults (Anblagan et al., 2018) and correlates more strongly negatively with recall performance than does hippocampal volume (den Heijer et al., 2012; Aribisala et al., 2014; Carlesimo et al., 2010; van Norden et al., 2012). Both macro- and microstructural properties of the hippocampus may contribute to developmental changes in successful recall of episodic content. In our previous life-span study, using an overlapping sample (Langnes et al., 2019), we found relationships between both hippocampal macro- and microstructure and recall scores over 30 min intervals, but these were dependent on the common influence from age.

In the present study, our first aim was to examine whether recall tested on average 10 days after encoding showed unique developmental trajectories that could not be accounted for by recall performance 30 min after encoding. In a previous cross-sectional study of 8–19 year olds, we found no developmental effects on visuo-constructive 1 week recall when 30 min recall performance was accounted for (Ostby et al., 2012), but performance on the extended retention interval recall test correlated with total hippocampal volume. In the present study, we used a larger sample ( $n = 650$  vs.  $n = 107$ ), including longitudinal observations (312 longitudinal examinations) and a wider age-range (4.1–24.8 vs. 8–19 years). This allowed us to assess developmental changes with higher sensitivity and superior statistical power. Both visuo-constructive and verbal recall tests were used. In addition, verbal and performance tests from Wechsler's intelligence batteries (Wechsler, 1999, 2008) were administered to allow us to test to what degree development of recall performance overlapped with general ability level.

Our second aim was to relate development of performance on recall tests using extended retention intervals to structural maturation of sub-regions of the hippocampus. Hippocampus was divided in an anterior (aHC) and a posterior (pHC) part according to established procedures (Poppenk et al., 2013). Macro-structural maturation was measured as regional volume and micro-structural maturation as regional mean diffusion (MD). We hypothesized that unique developmental trajectories would be seen for performance on recall tests using extended retention interval, and that performance would be positively related to regional hippocampal volume and negatively to regional hippocampal MD. Whether aHC or pHC would be more strongly related to recall performance is an open question, as previous results have not been consistent. Since connectivity studies have revealed distinct connections between the hippocampal regions and the neocortex, and there are few direct connections between aHC and pHC (Poppenk et al., 2013), it is conceivable that they may play different roles in the hippocampal-neocortical replay necessary for recall performance after extended time intervals. Previous studies have shown a role for prefrontal cortical regions in development of aspects of memory functions (see e.g. Shing et al., 2010; Ofen et al., 2007), and we also included an analysis testing the relationship between memory and thickness change in the lateral prefrontal cortex to allow comparisons with the hippocampal results.

## 2. Materials and methods

### 2.1. Sample

Participants were drawn from studies coordinated by the Research Group for Lifespan Changes in Brain and Cognition (LCBC [www.oslo.brains.no](http://www.oslo.brains.no)) (Fjell et al., 2015), approved by a Norwegian Regional Committee for Medical and Health Research Ethics. Written informed consent was obtained from all participants older than 12 years of age and from a parent/guardian of volunteers under 16 years of age. Oral informed consent was obtained from participants under 12 years of age. The full sample consisted of 650 healthy participants, 4.1–24.8 years of age (mean examination age, 10.4 years, 1<sup>st</sup> quartile = 6.7 years, 3<sup>rd</sup> quartile = 12.9) with a total of 914 MRI examinations and up to 832 recall tests sessions (832 observations for CVLT 30 min recall; 770 for CVLT extended retention interval; 666 for CFT 30 min; 602 for CFT extended retention interval). Participants were followed for up to 4 time points with MRI, for a maximum period of 8.9 years since baseline (mean interval between visits = 1.8 years, mean total follow up time since baseline for the longitudinal examinations = 1.9 years). Adult participants (> 20 years) were screened using a standardized health interview prior to inclusion in the study (see Langnes et al., 2019). Participants with a history of self- or parent-reported neurological or psychiatric conditions, including clinically significant stroke, serious head injury, untreated hypertension, diabetes, and use of psychoactive drugs within the last two years, were excluded. Further, participants reporting worries concerning their cognitive status, including memory function, were excluded.

### 2.2. Testing of recall performance

The California Verbal learning Test (CVLT) (Delis et al., 2000) and the Rey-Osterrieth Complex Figure (CFT) test (Meyers and Meyers, 1995) were used to assess recall. The 30 min retention interval recall conditions were used for both tests.

The learning part of CVLT consists of oral presentation of 16 words, at one second intervals, in four semantic categories (words are not presented ordered by category), and the whole list is presented five times with a free recall trial after each presentation. After five presentations, the free recall trials are repeated 5 and 30 min later. In addition, an additional extended retention interval recall condition was administered after a mean interval of 9.9 days (1<sup>st</sup> quartile = 7 days, 3<sup>rd</sup> quartile = 10 days).

For visual recall, CFT uses a novel, complex design which participants are asked to copy and then reproduce from memory after 30 min. The participants were presented with a picture of a geometrical figure on an A4 sheet of paper and were asked to draw the figure as similar as possible. After approximately 30 min, during which time the participants completed other tasks with mainly verbal material, they were asked to draw the figure again without the original picture in front of them. The scoring system divides the figure into 18 subunits and awards 2 points for each correct and correctly placed unit, 1 point for an inaccurately drawn or incorrectly placed unit, and a 1/2 point for a unit that is recognizable but both inaccurate and inaccurately placed in the drawing. This results in a maximum score of 36 points for each drawing. As for CVLT, an extended retention interval recall condition was administered after a mean interval of 10.2 days (1<sup>st</sup> quartile = 7 days, 3<sup>rd</sup> quartile = 10 days).

### 2.3. General ability testing

General cognitive abilities were assessed by Wechsler's Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999) for participants aged 6.5–89 years of age, while scores for corresponding subtests (Verbal: Vocabulary, Similarities; Visuo-constructive: Block design and Matrices) from the Wechsler Preschool and Primary Scale of intelligence – III

(WPPSI-III)(Wechsler, 2008) were used for the youngest participants (< 6.5 years).

#### 2.4. MRI acquisition and cross-scanner validation

858 scans were obtained from 1.5 T Avanto (12 channel head coil) and 56 from 3 T Skyra (20 channel head coil). The following sequences were used:

Avanto T1-weighted: 2 repeated 3D T1-weighted magnetization prepared rapid gradient echo (MPRAGE): TR/TE/TI = 2400 ms/3.61 ms/ 1000 ms, FA = 8°, acquisition matrix 192 × 192, FOV = 240 × 240 mm, 160 sagittal slices with voxel sizes 1.25 × 1.25 × 1.2 mm. For most children 4–9 years old, iPAT was used, acquiring multiple T1 scans within a short scan time, enabling us to discard scans with residual movement and average the scans with sufficient quality.

Avanto DTI: 32 directions, TR = 8200 ms, TE = 81 ms, b-value = 700 s/mm<sup>2</sup>, voxel size = 2.0 × 2.0 × 2.0 mm, field of view = 128, matrix size = 128 × 128 × 64, number of b0 images = 5, GRAPPA acceleration factor = 2.

Skyra T1-weighted: 176 sagittal oriented slices were obtained using a turbo field echo pulse sequence (TR = 2300 ms, TE = 2.98 ms, flip angle = 8°, voxel size = 1 × 1 × 1 mm, FOV = 256 × 256 mm). For the youngest children, integrated parallel acquisition techniques (iPAT) was used, acquiring multiple T1 scans within a short scan time, enabling us to discard scans with residual movement and average the scans with sufficient quality.

Skyra DTI: A single-shot twice-refocused spin-echo echo planar imaging (EPI) with 64 directions: TR = 9300 ms, TE = 87 ms, b-value = 1000s/mm<sup>2</sup>, voxel size = 2.0 × 2.0 × 2.0 mm, slice spacing = 2.6 mm, FOV = 256, matrix size = 128 × 130 × 70, 1 non-diffusion-weighted (b = 0) image. A b0-weighted image was acquired with the reverse phase encoding.

Since different scanners will yield different volumetric and MD values, 180 participants evenly distributed across a wide age range were scanned on the 1.5 T (Avanto) scanner and the 3 T (Skyra) scanner on the same day, to allow us to directly assess the influence of scanner. These data are previously published (Langnes et al., 2019), showing that the different scanners yielded significant differences in absolute MD and volume. The correlations between scanners were good, however,  $r = .85$  (anterior) and  $.88$  (posterior) for volume and  $.71$  (anterior) and  $.73$  (posterior) for MD. The high rank-order coherence between scanners suggested that inclusion of scanner as a covariate in the analyses efficiently would remove most of the variance between participants due to different scanners.

#### 2.5. MRI preprocessing - morphometry

T1-weighted scans were run through FreeSurfer 6.0 (<https://surfer.nmr.mgh.harvard.edu/>). FreeSurfer is an almost fully automated processing tool (Fischl et al., 1999a, 2002; Dale et al., 1999; Fischl et al., 1999b), and manual editing was not performed to avoid introducing errors. For the children, the issue of movement is especially important, as it could potentially induce bias in the images (Reuter et al., 2015). Rather, all scans were manually rated for movement on a 1–4 scale, and only scans with ratings 1 and 2 (no visible or only very minor possible signs of movement) were included in the analyses, reducing the risk of movement affecting the results. Also, all reconstructed surfaces were inspected, and discarded if they did not pass internal quality control. The hippocampus was initially segmented as part of the FreeSurfer subcortical stream (Fischl et al., 2002) before being divided in aHC and pHC (see below). 90 scans were discarded due to low quality due mainly to excessive motion, technical issues during acquisition, incomplete protocols (e.g. lacking DTI data) or reconstruction or segmentation issues, reducing the number of scans in the analyses to 824 for volume and mean diffusion and 877 for cortical thickness (see below).

#### 2.6. MRI preprocessing – DTI

DTI scans were processed with FMRIB's Diffusion Toolbox ([fsl.fmrib.ox.ac.uk/fsl/fslwiki/](http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/)) (Jenkinson et al., 2012; Smith et al., 2004). B0 images were also collected with reversed phase-encode blips, resulting in pairs of images with distortions going in opposite directions. From these pairs we estimated the susceptibility-induced off-resonance field using a method similar to what is described in (Andersson et al., 2003) as implemented in FSL (Smith et al., 2004). We then applied the estimate of the susceptibility induced off-resonance field with the eddy tool (Andersson and Sotiropoulos, 2016), which was also used to correct eddy-current induced distortions and subject head movement, align all images to the first image in the series and rotate the bvecs in accordance with the image alignments performed in the previous steps (Jenkinson et al., 2002; Leemans and Jones, 2009).

#### 2.7. Hippocampal anterior-posterior segmentation

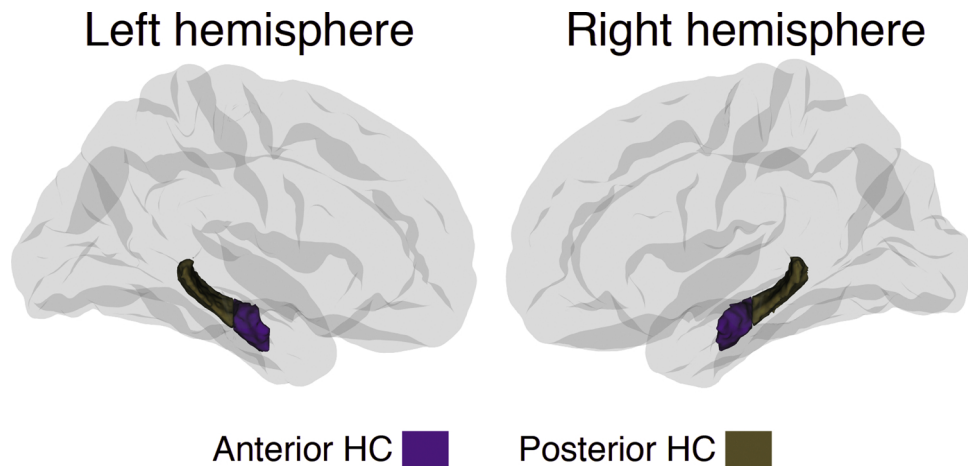
Moving anteriorly through the coronal planes of an MNI-resampled human brain,  $y = -21$  corresponds to the appearance of the uncus of the parahippocampal gyrus. In line with recent recommendations for long-axis segmentation of the hippocampus in human neuroimaging (Poppenk et al., 2013), we labeled hippocampal voxels at or anterior to this landmark as anterior HC while voxels posterior to the uncus apex were labeled as posterior HC. Specifically, for each participant, all diffusion voxels for which more than 50 % of the underlying anatomical voxels were labeled as hippocampus by FreeSurfer (Fischl et al., 2002) were considered representations of the hippocampus. While keeping the data in native subject space, we next established hippocampal voxels' locations relative to MNI  $y = -21$  by calculating the inverse of the MNI-transformation parameters for a given subject's brain and projecting the back-transformed coronal plane corresponding to MNI  $y = -21$  to diffusion native space. All reported diffusion measures thus represent averages from hippocampal sub-regions established in native space. An illustration of this segmentation is shown in Fig. 1. Segmentation results for examples of participants below 6 years are shown in Supplemental Information.

#### 2.8. Lateral prefrontal cortex parcellation

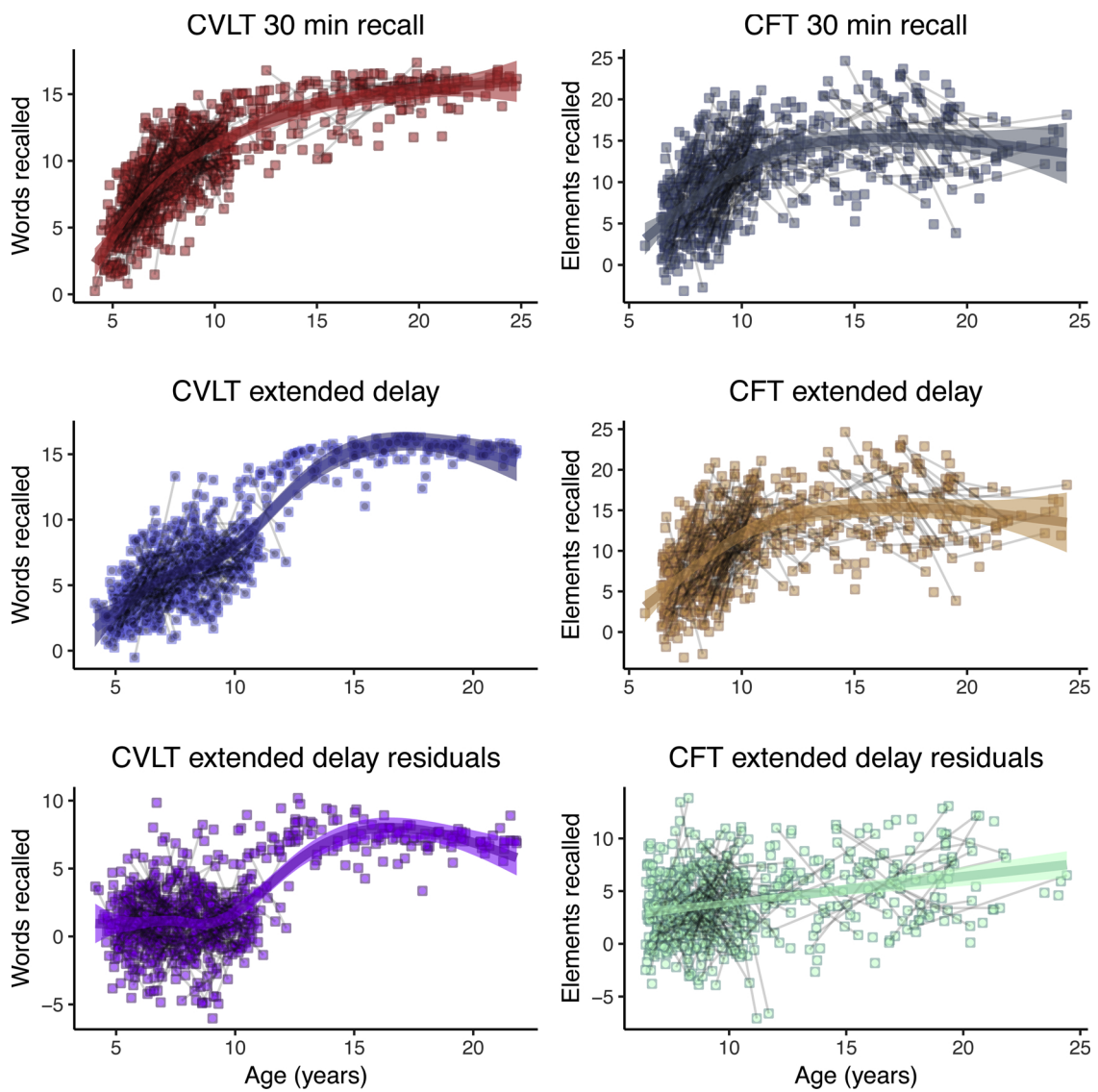
The lateral prefrontal cortex is often implied in performance of memory-related tasks, and thus a region of interest was defined from pars triangularis, pars orbitalis, pars opercularis and rostral middle frontal cortex across hemispheres, weighted by the size of each region, in the Desikan-Killiany parcellation scheme implemented in FreeSurfer (Desikan et al., 2006).

#### 2.9. Statistical analyses

Analyses were run in R (<https://www.r-project.org/>) using Rstudio ([www.rstudio.com](http://www.rstudio.com)) integrated development environment. Generalized Additive Mixed Models (GAMM) using the package "mgcv" (Wood, 2006) were used to derive age-functions. Different models were run with the recall scores from CVLT and CFT in turn as dependent variables. We included a smooth term for age, random effect for subject, and sex as covariate of no interest. Subject time-point was included as an additional covariate in all analyses of recall performance to control for practice effects of the scores. To test for unique developmental trajectories of recall performance after extended retention intervals, recall performance after 30 min was used as an additional covariate. If the relationship between recall performance after the long retention interval and age was still significant, this was taken as evidence of a unique developmental effect on the extended retention interval recall performance. The GAMM functions with recall performance after extended retention intervals as dependent and age as a smooth predictor, with 30 min recall performance, sex, retention interval and time point as



**Fig. 1.** Hippocampal sub-regions. Hippocampus was segmented in an anterior and a posterior part according to established procedures.



**Fig. 2.** Developmental trajectories for memory. Development of CVLT (left column) and CFT (right column) recall performance. The plots in the bottom row show how performance in the 10 days retention interval recall condition improves when recall performance on the 30 min retention interval condition is accounted for.



covariates were repeated with raw scores from WASI/ WPSI as additional covariates to test for the specificity of the age-recall performance relationships. Then each hippocampal sub-region and modality (volume, MD) was used as dependent variables in separate analyses. Scanner was used as an additional covariate of no interest in all analyses of hippocampal sub-regions. To assess the relationship between hippocampal sub-regions and recall performance, age and hippocampal sub-regions were included in the same models, with the same covariates as above.

In all the models, the smoothness of the age-curve is estimated as part of the model fit, and the resulting effective degrees of freedom (edf) was taken as a measure of deviation from linearity. The p-values associated with the smooth terms are only approximate, as they are based on the assumption that a penalized fit is equal to an unpenalized fit with the same edf, and do not take into account uncertainty associated with the smoothing parameter estimation. The major advantage of GAMM in the present setting is that relationships of any degree of complexity can be modelled without specification of the basic shape of the relationship, and GAMM is thus especially well-suited to map life-span trajectories of neurocognitive variables which can be assumed to be non-linear and where the basic form of the curve is not known (Fjell et al., 2010).

Since not all information was available for all participants, the number of observations that were used in the main analyses is presented.

### 3. Results

#### 3.1. Age-relationships

Trajectories of recall performance across age are presented in Fig. 2. 30 min recall scores showed the expected sharp development through childhood for both CVLT ( $F = 158.6$ ,  $edf = 5.0$ ,  $p < 2e^{-16}$ ,  $n = 832$ ) and CFT ( $F = 56.6$ ,  $edf = 43.9$ ,  $p < 2e^{-16}$ ,  $n = 666$ ). While CVLT score increased throughout the age-span, CFT peaked in mid adolescence. Next, the same analyses were run for recall performance tested after the extended retention interval. Both CVLT ( $F = 156.9$ ,  $edf = 4.9$ ,  $p < 2e^{-16}$ ,  $n = 752$ ) and CFT ( $F = 57.9$ ,  $edf = 3.4$ ,  $p = 2e^{-16}$ ) showed rapid increases through childhood and adolescence. For CVLT, however, the developmental trajectory for performance in the extended retention interval condition showed an earlier and steeper peak than in the 30 min condition, suggesting developmental differences. Thus, the GAMMs were re-run with 30 min performance as an additional covariate. These analyses showed clear effects of age on recall performance after the extended retention interval, independently of 30 min recall performance, both for CVLT ( $F = 64.7$ ,  $edf = 5.1$ ,  $p < 2e^{-16}$ ) and CFT ( $F = 31.6$ ,  $edf = 1$ ,  $p = 2.9e^{-08}$ ). For CFT, the residual age-trajectory of the extended retention interval recall performance was linear and positive when 30 min recall performance was accounted for. For CVLT, there was no additional developmental improvements over and above 30 min recall until about 10 years, from which sharp positive development was seen until the last part of adolescence.

#### 3.2. Influence of development of general cognitive abilities

We were interested in testing whether the developmental effects on recall performance after the extended retention interval were related to development of general verbal (for CVLT) and performance/ visuospatial (for CFT) abilities. Thus, we re-ran the GAMM functions with CVLT extended retention interval score as dependent and age as a smooth predictor, with 30 min recall, sex, retention interval and time point as covariates. In addition, we added similarities and vocabulary raw scores from the Wechsler tests. The relationship between age and CVLT recall performance after the extended retention interval was still significant ( $F = 27.3$ ,  $edf = 6.8$ ,  $p < 2e^{-16}$ ,  $n = 737$ ), while neither of the Wechsler tests were significantly related to age (both  $p$ 's  $> .15$ ). Since the Wechsler variables are highly correlated (similarities – vocabulary  $r = .74$ ; matrix reasoning – block design  $r = .71$ ), we re-ran the

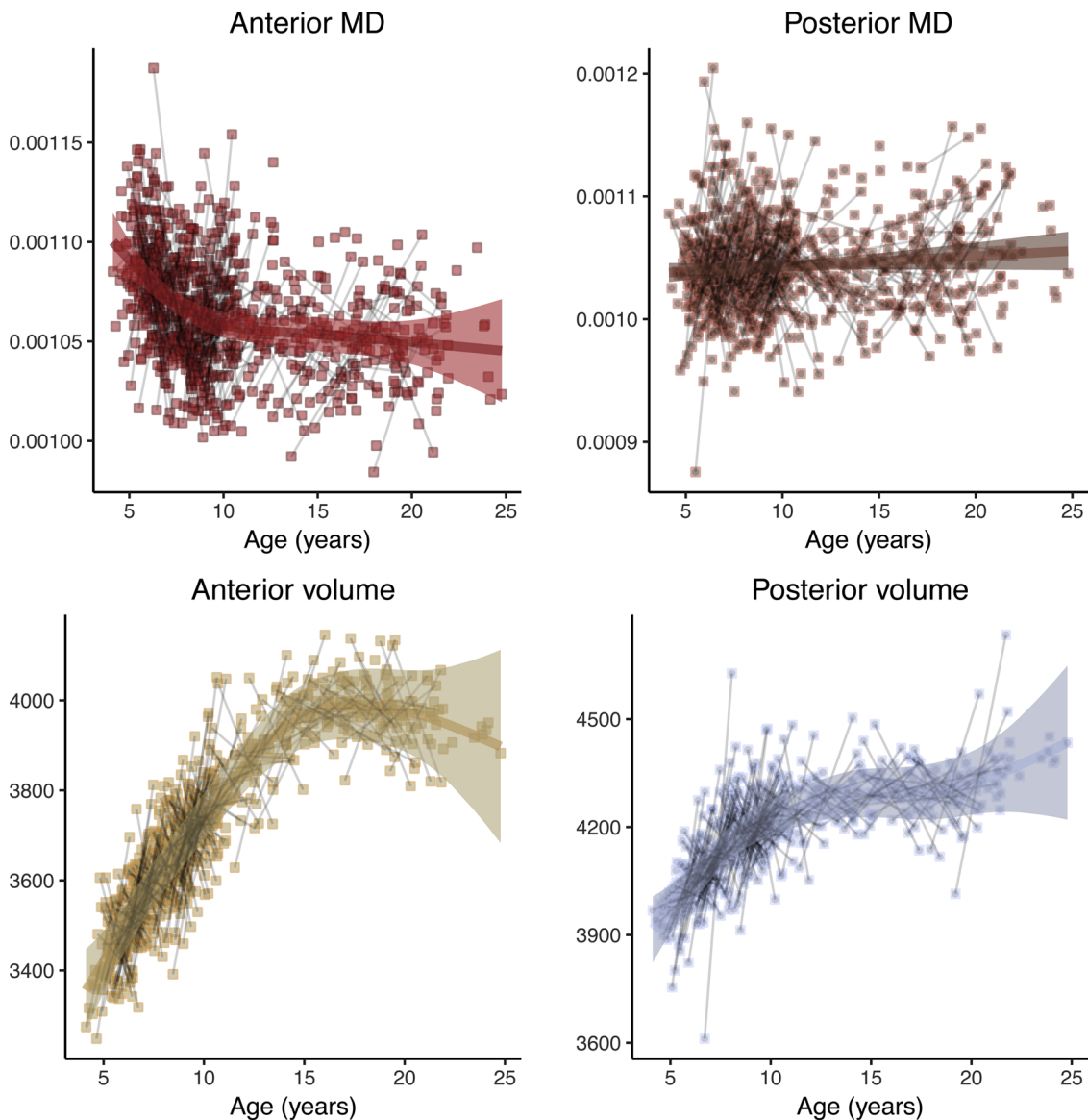
analysis with vocabulary and similarities in separate models. Now vocabulary contributed significantly and positively, although modestly, to recall performance after the extended retention interval (vocabulary:  $t = 2.2$ ,  $p = .03$ ,  $n = 741$ ; similarities:  $t = 1.8$ ,  $p = .076$ ,  $n = 737$ ). Formally testing whether verbal abilities significantly affected the developmental trajectory of CVLT recall after the extended retention interval, we calculated the residual age-function with and without the WASI variables as covariates. We then tested whether the derivatives of the models differed at any age. This was not the case, which implies that verbal ability levels do not significantly affect the age-trajectory of recall performance after extended retention intervals (see Fig. 6).

The same type of analysis was run with CFT score after the extended retention interval as dependent and the WASI measures of block design and matrix reasoning as covariates. Inclusion of these rendered the effect of age not significant ( $F = 0.06$ ,  $edf = 1$ ,  $p = .81$ ,  $n = 591$ ). When tested in separate models, both block design ( $t = 4.9$ ,  $p = p < 1.25e^{-6}$ ,  $n = 591$ ) and matrix reasoning ( $t = 3.2$ ,  $p = .001$ ,  $n = 599$ ) were significantly and positively related to performance in the extended retention interval condition. As for CVLT, we performed a formal test of the effects of the WASI scores on the developmental trajectory of CFT recall performance after the extended retention interval. Testing the derivatives of the age-function with vs. without the WASI tests as covariates revealed a significant influence on the age-trajectory (see Fig. 6).

#### 3.3. Relationships to hippocampal volume and macrostructure

Developmental trajectories for volume and MD for hippocampal sub-regions are shown in Fig. 3. These are included as background information, since the current data have previously been used in a separate publication on life-span trajectories of hippocampal sub-region structure and the relationship to recall after the conventionally used 30 min interval (Langnes et al., 2019), in contrast to the present paper's focus on recall after the extended retention interval. However, as the hippocampus-age relationships have been previously investigated in the larger sample, the p-values should be interpreted with caution. As expected, hippocampal anterior ( $F = 38.5$ ,  $edf = 3.7$ ,  $p < 2e^{-16}$ ,  $n = 824$  for all analyses) and posterior ( $F = 16.4$ ,  $edf = 3.8$ ,  $p = 4.8e^{-12}$ ) volume increased early in development, but either peaked towards the end of adolescence (anterior) or flattened out showing only modest changes from early teen years (posterior). For MD, the relationships differed substantially in the anterior vs. posterior regions. Anterior MD was reduced in childhood, and showed little further change during adolescence ( $F = 13.9$ ,  $edf = 2.9$ ,  $p = 9.3e^{-09}$ ). In contrast, posterior MD did not show any significant developmental effect ( $F = 2.7$ ,  $edf = 1$ ,  $p = .10$ ). These results demonstrate the volume and MD show different developmental trajectories, and that the anterior and posterior region of the hippocampus differ in development.

Finally, we tested the relationship between recall performance and hippocampal sub-regional volume or microstructure, controlling for age, sex, subject time point and scanner (number of observations for each analysis: CVLT 30 min  $n = 715$ ; CVLT extended retention interval  $n = 663$ ; CFT 30 min  $n = 615$ ; CFT extended retention interval  $n = 555$ ). The results are presented in Table 1 and Fig. 4 and illustrated with scatterplots in Fig. 5. Four relationships were found at an uncorrected  $\alpha$ -level of .05. Three of these survived adjusted Bonferroni corrections (16 tests, mean correlation between the dependent variables  $r = 0.565$ , critical corrected  $\alpha = 0.015$ ). Anterior hippocampal MD was related to CVLT performance after the extended retention interval as well as CFT 30 min recall, in both cases with lower MD being associated with higher recall score. The relationship to CVLT score was specific to recall performance after the extended retention interval, as it was not found for the 30 min condition, and still was significant when adding 30 min recall performance as an additional covariate ( $p = .014$ ). Although the relationship between CVLT extended retention interval recall and anterior volume did not survive corrections, we ran an additional GAMM including both anterior volume and MD simultaneously with the same covariates. The



**Fig. 3.** Developmental trajectories for hippocampus. Structural maturation of hippocampal sub-regions. Top row shows microstructure (mean diffusion), bottom row shows volume.

relationships to CVLT extended interval recall score were not weakened (anterior volume:  $F = 4.69$ ,  $p = .031$ ; anterior MD:  $F = 7.26$ ,  $p = .007$ ), suggesting that volume and microstructure are independently related to recall after extended retention intervals. In addition, posterior volume correlated positively with CFT recall after the extended interval. Since the relationships between each of the dependent variables in general were modest, we did not proceed to run additional models including all sub-regions/ modalities as simultaneous predictors.

### 3.4. Relationships to lateral prefrontal thickness

Similar GAMMs were run for thickness in the lateral prefrontal cortex. Significant negative relationships at  $p < .05$  uncorrected were found with CFT 30 min recall ( $F = 5.60$ ,  $p = .018$ ) and with recall after the extended retention interval ( $F = 3.91$ ,  $p = .049$ ), as well as a trend for CVLT 30 min recall ( $F = 3.20$ ,  $p = .074$ ) (see Fig. 5). Taking the mean correlation between the memory tests into account (mean  $r = .565$ ), the relationship between CFT 30 min recall and lateral prefrontal thickness would survive Bonferroni correction (critical  $\alpha = .027$ ). Since anterior MD also was a significant predictor of CFT 30 min recall, we ran

a GAMM including both as predictors with the same covariates. Both prefrontal thickness and anterior hippocampal MD were significant predictors in this model (anterior MD:  $F = 5.46$ ,  $p = .007$ ; lateral prefrontal thickness:  $F = 6.32$ ,  $p = .012$ ), showing that hippocampal microstructure and prefrontal thickness are independent predictors of CFT 30 min recall score.

## 4. Discussion

Three main conclusions can be drawn from the present study. First, there are distinct developmental trajectories for recall performance after extended retention intervals that cannot be explained by recall performance after shorter time intervals. Second, recall performance was modestly related to structural features of the hippocampus and the lateral prefrontal cortex, each explaining unique variance. Finally, visuo-constructive ability level significantly affected the developmental trajectory of CFT recall performance after extended retention intervals, explaining the major part of the age effect, while the developmental trajectory of CVLT recall performance after extended retention intervals was not affected by verbal ability level. The implications of the findings

**Table 1**  
Relationship between hippocampal sub-regions, lateral prefrontal cortex and memory.

	CVLT					
	30 min recall			Extended retention interval		
	edf	F	p <	edf	F	p <
Ant vol	1.0	3.1	.08	1.0	5.2	<b>.023</b>
Post vol	1.0	0.2	.69	1.0	0.3	.60
Ant MD	1.0	0.9	.37	1.0	7.7	<b>.006</b>
Post MD	1.0	0.4	.54	1.0	1.9	.17
LPFC	1.0	3.2	.07	1.0	2.0	.15

	CFT					
	30 min recall			Extended retention interval		
	edf	F	p <	edf	F	p <
Ant vol	1.0	2.5	.11	1.0	4.0	.11
Post vol	1.0	1.5	.22	1.0	6.6	<b>.011</b>
Ant MD	2.4	5.1	<b>.011</b>	1.0	0.3	.60
Post MD	1.0	0.1	.77	1.0	2.3	.13
LPFC	1.0	5.60	<b>.018</b>	1.0	3.9	<b>.049</b>

LPFC: Lateral prefrontal cortex thickness.

**Bold** indicate p < .05, Bonferroni corrected for multiple comparisons.

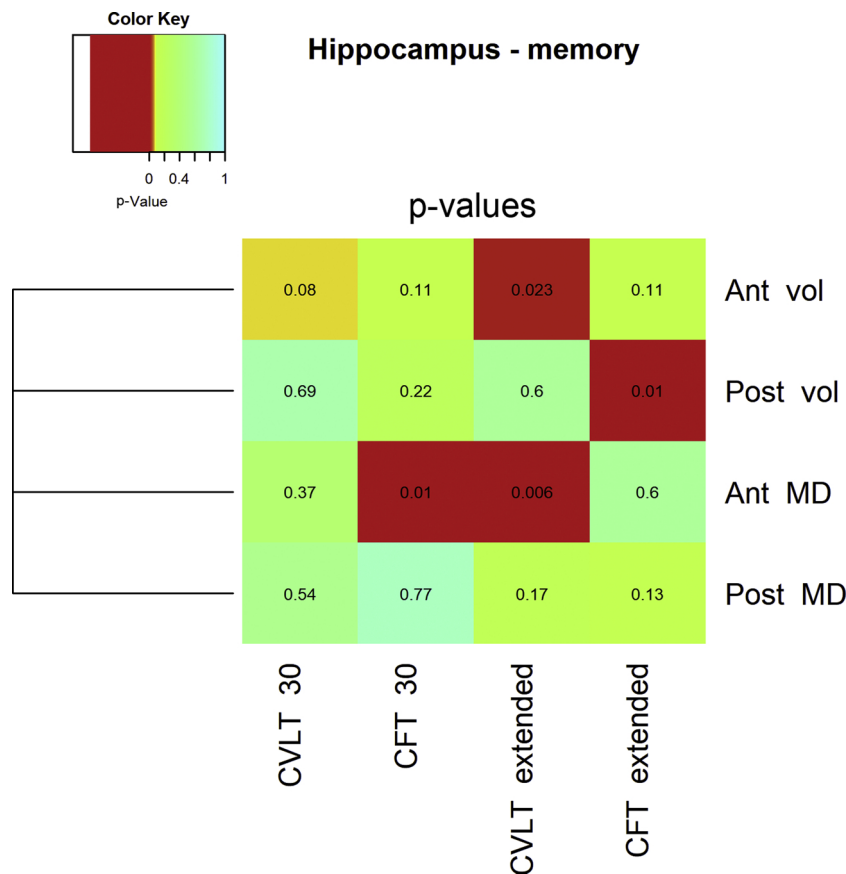
**Bold italics** indicate p < .05 uncorrected.

are discussed below.

4.1. Development of long-term recall

Although a primary function of episodic memory is to keep information in an accessible form over prolonged intervals, long term

episodic recall is usually tested after an hour or less. This is based on a premise that the processes responsible for successful long-term recall can be evaluated after short time intervals. However, there is good evidence from neuropsychological (Mayes et al., 2003) and molecular studies (Kandel, 2012) that episodic information is stored through an initial process of rapid consolidation followed by a slower consolidation phase, and that these consolidation phases depend on fundamentally different processes in the brain. Hippocampal-neocortical replay is ongoing for extended time after encoding (Moscovitch et al., 2016), and both hippocampus and its structural and functional connections show substantial developmental changes through childhood (Ostby et al., 2009; Tamnes et al., 2013; Lebel and Deoni, 2018; Langnes et al., 2018a; Lebel and Beaulieu, 2011). Thus, it would not be surprising if these extended consolidation processes are affected by maturational events in childhood and adolescence. Developmental effects on the processes responsible for stabilization, maintenance and transformations of episodic content over longer time intervals would be expected to result in specific developmental effects on extended retention interval recall that cannot be accounted for by recall performance over shorter time. This was exactly what was found in the present study, for two very different recall tasks. There were unique developmental effects on both CFT and CVLT recall performance after extended retention intervals, over and above scores on the 30 min recall conditions. In adults, the phenomenon of accelerated long-term forgetting (ALF) has been used to refer to abnormal forgetting over hours to weeks despite normal encoding or initial consolidation (Elliott et al., 2014; Blake et al., 2000). This suggests that recall performance after longer time intervals is supported by partly separate brain processes than recall performance after 30 min, which is in line with the unique developmental effects observed for recall performance after extended retention intervals in the



**Fig. 4.** Memory and hippocampal subfields p-values.

The heat plot illustrates the statistical significance (uncorrected p-value) of the relationship between hippocampal structure (volume and mean diffusion) and memory performance. Age was included as a covariate in the analyses.

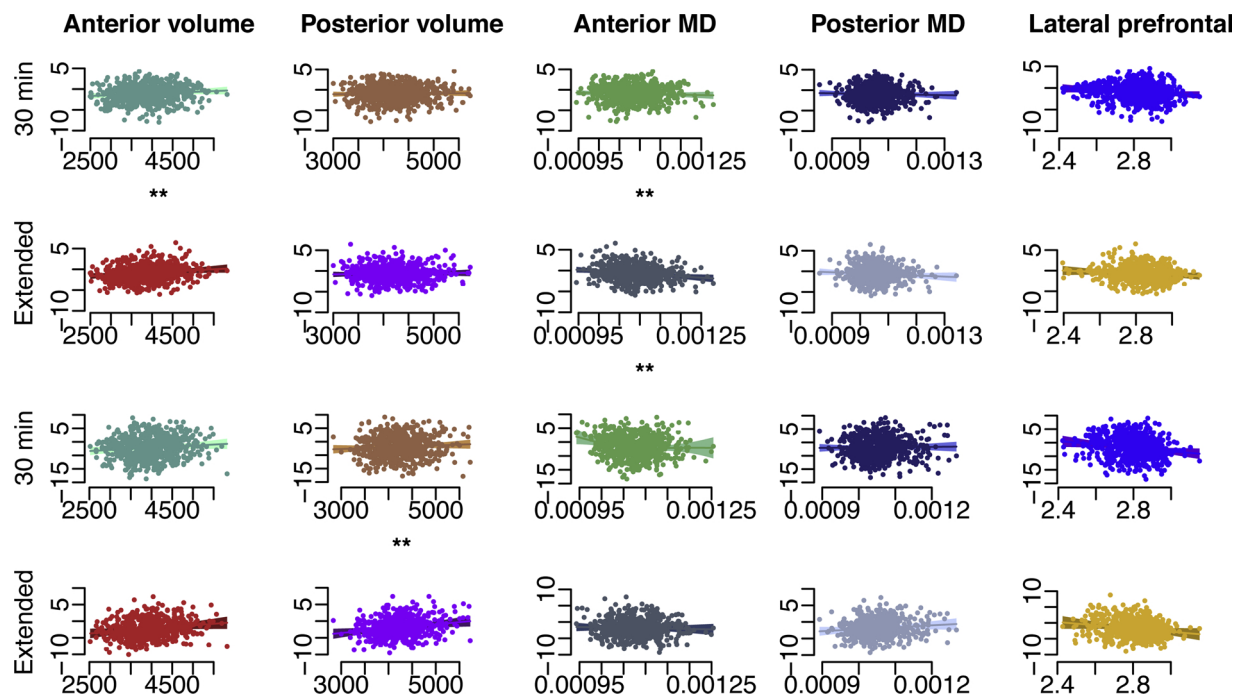


Fig. 5. Memory and hippocampal subfields relationships.

Scatterplots illustrating the relationship between hippocampus structure and memory performance. The two top rows show CVLT recall performance. The two bottom rows show CFT performance.

\*\*  $p < .05$

present study. Similarly, although the evidence is partly mixed, there seems to be increased forgetting over long time intervals also in the other end of the lifespan (Elliott et al., 2014).

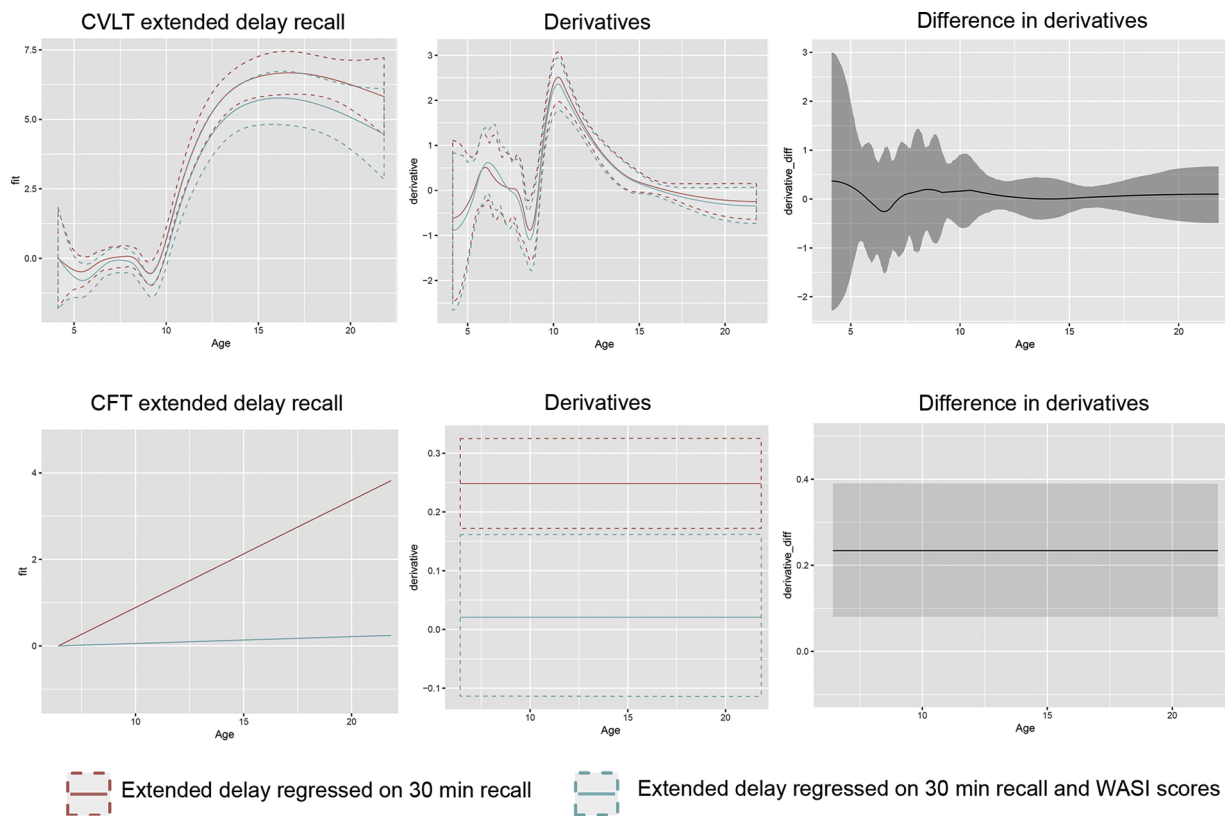
Importantly, however, it is a fallacy to conclude from the present results alone that consolidation processes are selectively changing through development. Brain activation studies have shown that hippocampal-neocortical connectivity during encoding can explain differences between recall performance over hours compared to weeks (Sneve et al., 2015). In one such study it was found that a critical level of activation of the hippocampus during encoding was necessary for source recall performance for both hours and weeks, while strong levels of connectivity between hippocampus and perceptual and self-referential default mode networks during encoding were necessary for establishment of durable source memories (Sneve et al., 2015). This is relevant for the interpretation of the present results, because these effects were observed at the stage of encoding, in principle independently of extended consolidation processes. It is likely that the increase in hippocampal-cortical connectivity seen during encoding also affected consolidation processes over extended intervals, but this is a speculation. Similarly, we do not know to what degree the unique developmental trajectories for recall performance over extended time intervals is due to differences in encoding or consolidation, or, most likely, a combination of them. In any case, the present results, in combination with the known maturational changes in structures and connections of known importance for episodic memory consolidation, makes a case for improved efficiency of consolidation processes during childhood and adolescence.

The present results also showed that although verbal ability level as measured by the WASI subtest vocabulary was modestly related to CVLT recall performance in the extended retention interval condition, the age-trajectory was minimally affected by controlling for scores on the verbal WASI tests. In contrast, development of CFT recall performance after the extended retention interval was highly influenced by scores on the matrix reasoning and block design tests, especially the latter. Controlling for performance on these two tests completely removed the

developmental effect on CFT recall performance after the extended retention interval. This is in line with previous observations that the copy score on the CFT improves significantly in development and is highly related to short- and long-term recall (Ostby et al., 2012), but extends these by showing that performance on independent ability tests account for the developmental improvements in recall abilities. The strong effects of visuo-constructive abilities on CFT recall performance after the extended retention interval may be an instance of statistical collinearity, and the developmental effects on recall after extended retention intervals may be real but impossible to disentangle from developmental effects in other cognitive domains. However, the effects may also be caused by higher visuo-constructive abilities leading to more efficient or elaborative encoding strategies due to better understanding and conceptualization of the material, again causing superior organization and hence improved recall (Chase and Simon, 1973). Large effects of understanding of the material to be learned on encoding efficiency is a well-established result in cognitive psychology (Sala and Gobet, 2017), and this phenomenon is likely to at least partly explain the observed effect of visuo-constructive ability level on the developmental trajectory of CFT recall performance after the extended retention interval. One reason for the effect of ability level being more important for CFT than CVLT recall may be that mental organization and comprehension of the CFT figure is more demanding than the CVLT words, and differences in general ability levels thus may influence CFT recall to a larger extent than CVLT recall. Still, good categorization of the words in CVLT also represents a major mnemonic tool, which requires a certain verbal ability level. However, the effects of verbal ability level on verbal performance in the extended retention interval condition were small beyond what was already present at the 30 min interval.

There were also interesting differences in the age-trajectories for CFT vs. CVLT recall scores for the extended retention interval when performance at the 30 min interval was accounted for. CVLT showed only slight increases until about 10 years, after which the residual scores increased sharply. In contrast, the residual scores for CFT increased linearly through the age-range. The two tests represent fundamentally





**Fig. 6.** Effects of general ability levels. We tested how WASI scores affected the age-relationship of recall performance after the extended retention interval. Top row shows the result for CVLT. Bottom row shows the results for CFT. In the first column, the age-trajectory without ability level controlled for is shown in red, and the age-trajectory with ability level controlled for in blue-cyan. In the middle row, the respective derivatives of the age-curves are shown. In the right column, the differences in derivatives between the red and the blue-cyan curves are plotted. When the confidence interval of these differences does not include zero, this means that the effect of ability level on the age-trajectory is significant. This is the case for CFT recall performance after the extended retention interval (bottom right corner), but not CVLT recall performance after the extended retention interval (top right corner). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

different task demands, such as intentional vs. incidental encoding, clear semantic vs. more abstract – although not completely – material, auditory vs. visual presentation, fixed presentation rate vs. self-paced, and different challenges in categorizing the material. Thus, it is not possible to point to one single mechanism that can explain the different developmental trajectories for the two tests. It can be speculated that for instance maturation of strategic encoding and retrieval strategies accelerates at the age we see the increase in CVLT residual scores, and that these influence CVLT scores more than CFT due to the intentional nature of the task as well as the huge beneficial effects of successful categorization of the words. One framework that could be applied involves a distinction between a strategic component of memory supported by the frontal cortex and an associative component supported by the medial temporal lobe, especially the hippocampus (Shing et al., 2010). Prolonged development of the prefrontal cortex (Tamnes et al., 2013; Gogtay et al., 2004) could then contribute to increased efficiency in use of strategic components while performing the memory tasks. This would then assume that improvements in the strategic component affect CVLT performance over longer retention intervals more than CFT performance. As argued above, copy scores on the CFT, as well as score on the visuo-constructive Wechsler tests, are tightly connected to the developmental improvements in CFT, which may very well be caused by better organizational strategies. Further, lateral prefrontal thickness development was not significantly related to CVLT performance, but instead showed relationships to CFT, especially in the 30 min retention condition (see below).

#### 4.2. Effects of hippocampal structure and lateral prefrontal cortex

The major theories of long-term memory, such as the standard consolidation theory (Squire and Alvarez, 1995) and multiple trace theory (Nadel and Moscovitch, 1997), would predict that the efficiency of hippocampal processes affect consolidation over the time interval used in the extended retention interval conditions in present study. A structurally immature hippocampus would likely lead to less efficient re-activation of encoded memories and a less efficient hippocampal-cortical dialogue. In the framework of the multiple trace theory, this may be seen as less efficient establishment of multiple traces in the medial temporal lobe and the neocortex - and consequently steeper forgetting rates. The present finding of unique developmental effects for recall performance after the extended retention interval fits this hypothesis, further supported by the existence – although modest – of relationships with hippocampal sub-region structure. This interpretation is in concordance with an earlier cross-sectional study with a sample overlapping the present one, where CFT recall performance after an extended retention interval was related to hippocampal volume while performance over 30 min was related to prefrontal cortex thickness (Ostby et al., 2012). The latter finding was replicated in the current study, although relationships between CFT extended retention interval recall score and lateral prefrontal cortex thickness cannot be excluded based on the present result.

In adults, the relationship between hippocampal volume and recall performance is generally not robust (Van Petten, 2004), but may be stronger with longitudinal designs (Fjell et al., 2013; Gorbach et al., 2017; Vidal-Pineiro et al., 2018), in development (Tamnes et al., 2014)

and with recall tests spanning longer time intervals (Walhovd et al., 2004). While we could not find any developmental studies examining the relationship between hippocampal microstructure and recall performance in development, studies of adults suggest that MD may more closely than volume be associated with individual differences in recall of successful episodic content (den Heijer et al., 2012; Aribisala et al., 2014; Carlesimo et al., 2010; van Norden et al., 2012). In the present study, four hippocampus-recall relationships were identified at a nominal  $\alpha$ -level of .05, with three surviving adjusted Bonferroni corrections, taking into account the correlations between the different variables. However, inspecting the heat map in Fig. 4, it is difficult to make inferences about stronger relationships with recall for one of the structural hippocampal measures over the other.

Similarly, there were not obvious differences in the recall performance relationships for the hippocampal sub-regions. The role of aHC vs. pHC in successful recall of episodic information depends on the specific (Poppenk et al., 2013; Chase et al., 2015; Kühn and Gallinat, 2014) and general demands of the task, and it has been suggested that aHC is more involved in encoding and pHC in retrieval (Poppenk et al., 2013; Kühn and Gallinat, 2014; Lepage et al., 1998; Nadel et al., 2012), but see (Lee et al., 2017). This, however, likely depends on fundamental differences in the specific cognitive processes supported by the sub-regions, such as relational processing (Maass et al., 2014; Milivojevic et al., 2015), attention to perceptual aspects of the stimulus (Moscovitch et al., 2016; Poppenk et al., 2013), violation of narrative predictions (Milivojevic et al., 2015) and communication with different large-scale cortical networks (Kim, 2015). These different accounts are discussed in several comprehensive reviews (Poppenk et al., 2013; Lee et al., 2017) and meta-analyses (Kühn and Gallinat, 2014; Spaniol et al., 2009), showing the complexity of the relationships between the specific processes supported by hippocampal sub-regions. Also, involvement of hippocampal sub-regions in different recall tasks seems to be a matter of degree. For instance, one recent study observed the expected pattern that pHC was relatively more involved during retrieval than aHC, and that all sub-regions were active during encoding (Hrybowski et al., 2019), while a longitudinal developmental study found lower activation in the hippocampal tail than the body during retrieval (Selmecky et al., 2019).

Using only behavioral testing, we could not distinguish between encoding and retrieval effects on recall performance after the extended retention interval. We have previously found that children engaged pHC more than aHC, while aHC was more activated relative to pHC already in teenagers (Langnes et al., 2018b). The partly different structural developmental trajectories of the hippocampal sub-regions may have impact on more specific processes, which we were not able to detect with our behavioral task. Positive correlations between both aHC (Riggins et al., 2015) and pHC (DeMaster et al., 2014) have been reported in development, as in adults (Hackert et al., 2002; Poppenk and Moscovitch, 2011; Driscoll et al., 2003). The present results did not show consistent aHC-pHC differences. Rather, the effects varied as a function of retention interval, with tendencies for stronger relationships between hippocampal sub-regions and recall performance at the extended retention intervals than the 30 min interval, as discussed above.

#### 4.3. Limitations

There are multiple limitations of the current study. For the oldest participants, some approached the maximum score, which makes it possible that ceiling effects affected the results. Inspections of the scatterplots in Fig. 2 do not indicate that this is a serious issue, but this may still have contributed to reduce the observed age-effects, especially among the oldest adolescents. Among the youngest participants, some showed low scores, which may suggest possible floor effects in this age-range. As the scores increased rapidly with advancing age, however, this has likely not affected the age-trajectories substantially. We did not directly correct for selective attrition or learning effects, which may

impact the results (Josefsson et al., 2016, 2012; Nyberg et al., 2012). Instead, we included subject time-point as a covariate in all analyses, which should effectively control for the effects of taking the test multiple times. Another difficult issue when studying long-term recall is how to deal with differences in initial learning rate. In the present study, scores on the 30 min condition were regressed out in the models including long-term memory, yielding good statistical control for differences in initial performance level. This is a challenging issue, however, with multiple possible solutions with different strengths and weaknesses. Some have matched initial performance levels by e.g. multiple repetitions of items during encoding. However, no consensus has been reached regarding whether or not degree of initial learning affects rate of forgetting, and there is presently no agreement about how best to tackle this problem (see (Elliott et al., 2014) for an extensive discussion of these issues). Finally, the present study focused on the hippocampus, with additional analyses of a large region of the lateral prefrontal cortex. Obviously, brain structure - recall relationships could have been found in other brain regions not included in the present work.

## 5. Conclusion

CFT and CVLT recall performance after extended retention intervals showed unique developmental trajectories not accounted for by recall performance after a retention interval of 30 min. For CFT recall performance after the extended retention interval, these improvements during childhood and adolescence could be accounted for by general visual-performance ability level. This was not the case for CVLT recall performance after the extended retention interval, where the developmental trajectory was not affected by verbal ability. Finally, recall performance was modestly related to hippocampal and lateral prefrontal cortex structure. Experimental work, including functional brain activation studies, will be necessary to reveal to which degree development of the ability to recall information after retention interval spanning days in contrast to hours or less are caused by maturation of consolidation processes versus encoding-related processes. We believe task-related fMRI using memory tasks with different retention intervals may be crucial in further elucidating the neurocognitive foundation for development of long-term memory recall and consolidation processes in development.

## Funding

This work was funded by the Department of Psychology, University of Oslo (to K.B.W., A.M.F.), the Norwegian Research Council (to K.B.W., A.M.F.) and the project has received funding from the European Research Council's Starting/ Consolidator Grant schemes under grant agreements 283634, 725025 (to A.M.F.) and 313440 (to K.B.W.).

## Declaration of Competing Interest

None.

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.dcn.2019.100723>.

## References

- Anlagan, D., et al., 2018. Coupled changes in hippocampal structure and cognitive ability in later life. *Brain Behav.* 8 (2), e00838.
- Aribisala, B.S., et al., 2014. Quantitative multi-modal MRI of the Hippocampus and cognitive ability in community-dwelling older subjects. *Cortex* 53, 34–44.
- Bauer, P.J., 2015. Development of episodic and autobiographical memory: the importance of remembering forgetting. *Dev. Rev.* 38, 146–166.
- Bechmann, I., Nitsch, R., 2000. Involvement of non-neuronal cells in entorhinal-hippocampal reorganization following lesions. *Ann. N. Y. Acad. Sci.* 911, 192–206.

- Blake, R.V., et al., 2000. Accelerated forgetting in patients with epilepsy: evidence for an impairment in memory consolidation. *Brain* 123 (Pt 3), 472–483.
- Carlesimo, G.A., et al., 2010. Hippocampal mean diffusivity and memory in healthy elderly individuals: a cross-sectional study. *Neurology* 74 (3), 194–200.
- Chase, W.G., Simon, H.A., 1973. Perception in chess. *Cogn. Psychol.* 4 (4), 55–81.
- Chase, H.W., et al., 2015. Evidence for an anterior–posterior differentiation in the human hippocampal formation revealed by meta-analytic parcellation of fMRI coordinate maps: focus on the subiculum. *NeuroImage* 113, 44–60.
- Collin, S.H., Milivojevic, B., Doeller, C.F., 2015. Memory hierarchies map onto the hippocampal long axis in humans. *Nat. Neurosci.* 18 (11), 1562–1564.
- Dale, A.M., Fischl, B., Sereno, M.L., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 9 (2), 179–194.
- Daugherty, A.M., Flinn, R., Ofen, N., 2017. Hippocampal CA3-dentate gyrus volume uniquely linked to improvement in associative memory from childhood to adulthood. *Neuroimage* 153, 75–85.
- Delis, D.C., et al., 2000. California Verbal Learning Test - Second Edition (CVLT - II). The Psychological Corporation, San Antonio, TX.
- DeMaster, D., et al., 2014. Structural development of the hippocampus and episodic memory: developmental differences along the anterior/posterior axis. *Cereb. Cortex* 24 (11), 3036–3045.
- den Heijer, T., et al., 2012. Structural and diffusion MRI measures of the hippocampus and memory performance. *Neuroimage* 63 (4), 1782–1789.
- Desikan, R.S., et al., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31 (3), 968–980.
- Driscoll, L., et al., 2003. The aging hippocampus: cognitive, biochemical and structural findings. *Cereb. Cortex* 13 (12), 1344–1351.
- Elliott, G., Isaac, C.L., Muhlert, N., 2014. Measuring forgetting: a critical review of accelerated long-term forgetting studies. *Cortex* 54, 16–32.
- Fischl, B., et al., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum. Brain Mapp.* 8 (4), 272–284.
- Fischl, B., Sereno, M.L., Dale, A.M., 1999. Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9 (2), 195–207.
- Fischl, B., et al., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33 (3), 341–355.
- Fjell, A.M., et al., 2013. Brain changes in older adults at very low risk for Alzheimer's disease. *J. Neurosci.* 33 (19), 8237–8242.
- Fjell, A.M., et al., 2015. Development and aging of cortical thickness correspond to genetic organization patterns. *Proc. Natl. Acad. Sci. U. S. A.* 112 (50), 15462–15467.
- Geib, B.R., et al., 2017. Hippocampal contributions to the large-scale episodic memory network predict vivid visual memories. *Cereb. Cortex* 27 (1), 680–693.
- Gogtay, N., et al., 2004. Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U. S. A.* 101 (21), 8174–8179.
- Gogtay, N., et al., 2006. Dynamic mapping of normal human hippocampal development. *Hippocampus* 16 (8), 664–672.
- Goncalves, J.T., Schafer, S.T., Gage, F.H., 2016. Adult neurogenesis in the Hippocampus: from stem cells to behavior. *Cell* 167 (4), 897–914.
- Gorbach, T., et al., 2017. Longitudinal association between hippocampus atrophy and episodic-memory decline. *Neurobiol. Aging* 51, 167–176.
- Hackert, V.H., et al., 2002. Hippocampal head size associated with verbal memory performance in nondemented elderly. *Neuroimage* 17 (3), 1365–1372.
- Hrybouski, S., et al., 2019. Involvement of hippocampal subfields and anterior-posterior subregions in encoding and retrieval of item, spatial, and associative memories: longitudinal versus transverse axis. *Neuroimage* 191, 568–586.
- Jenkinson, M., et al., 2012. FSL. *Neuroimage* 62 (2), 782–790.
- Josefsson, M., et al., 2012. Genetic and lifestyle predictors of 15-year longitudinal change in episodic memory. *J. Am. Geriatr. Soc.* 60 (12), 2308–2312.
- Josefsson, M., et al., 2016. Causal inference with longitudinal outcomes and non-ignorable drop-out: estimating the effect of living alone on cognitive decline. *J. R. Stat. Soc. Ser. C Appl. Stat.* 65 (1), 131–144.
- Kandel, E.R., 2012. The molecular biology of memory: cAMP, PKA, CRE, CREB-1, CREB-2, and CPEB. *Mol. Brain* 5, 14.
- Kantor, D.B., Kolodkin, A.L., 2003. Curbing the excesses of youth: molecular insights into axonal pruning. *Neuron* 38 (6), 849–852.
- Kim, H., 2015. Encoding and retrieval along the long axis of the hippocampus and their relationships with dorsal attention and default mode networks: the HERNET model. *Hippocampus* 25 (4), 500–510.
- Kühn, S., Gallinat, J., 2014. Segregating cognitive functions within hippocampal formation: a quantitative meta-analysis on spatial navigation and episodic memory. *Hum. Brain Mapp.* 35 (4), 1129–1142.
- Langnes, E., et al., 2018. Development and decline of the hippocampal long-axis specialization and differentiation during encoding and retrieval of episodic memories. *Cereb. Cortex*.
- Langnes, E., et al., 2018. Development and decline of the hippocampal long-axis specialization and differentiation during encoding and retrieval of episodic memories. *bioRxiv*. <https://doi.org/10.1101/323097>.
- Langnes, E., et al., 2019. Lifespan trajectories and relationships to memory of the macro- and microstructure of the anterior and posterior hippocampus - a longitudinal multimodal imaging study. *bioRxiv*.
- Lebel, C., Beaulieu, C., 2011. Longitudinal development of human brain wiring continues from childhood into adulthood. *J. Neurosci.* 31 (30), 10937–10947.
- Lebel, C., Deoni, S., 2018. The development of brain white matter microstructure. *Neuroimage* 182, 207–218.
- Lee, J.K., J. E.G. Ghetti, S., 2017. Hippocampal development: structure, function and implications. In: Hannula, D.E., Duff, M.C. (Eds.), *The Hippocampus from Cells to Systems*. Springer, pp. 141–166.
- Lepage, M., Habib, R., Tulving, E., 1998. Hippocampal PET activations of memory encoding and retrieval: the HIPER model. *Hippocampus* 8 (4), 313–322.
- Lester, A.W., et al., 2017. The aging navigational system. *Neuron* 95 (5), 1019–1035.
- Lin, M., et al., 2013. Developmental changes in hippocampal shape among preadolescent children. *Int. J. Dev. Neurosci.* 31 (7), 473–481.
- Maass, A., et al., 2014. Laminar activity in the hippocampus and entorhinal cortex related to novelty and episodic encoding. *Nat. Commun.* 5, 5547.
- Mayes, A.R., et al., 2003. Long-term amnesia: a review and detailed illustrative case study. *Cortex* 39 (4–5), 567–603.
- Meyers, J.E., Meyers, K.R., 1995. Rey Complex figure Test and Recognition Trial. Professional Manual. Psychological Assessment Resources, Inc., Lutz (FL).
- Milivojevic, B., Vicente-Grabovetsky, A., Doeller, C.F., 2015. Insight reconfigures hippocampal-prefrontal memories. *Curr. Biol.* 25 (7), 821–830.
- Moscovitch, M., et al., 2016. Episodic Memory and Beyond: The Hippocampus and Neocortex in Transformation Recollective qualities modulate hippocampal activation during autobiographical memory retrieval. *Annu. Rev. Psychol.* 67 (1), 105–134.
- Nadel, L., Moscovitch, M., 1997. Memory consolidation, retrograde amnesia and the hippocampal complex. *Curr. Opin. Neurobiol.* 7 (2), 217–227.
- Nadel, L., Hoescheidt, S., Ryan, L.R., 2012. Spatial cognition and the Hippocampus: the anterior–posterior axis. *J. Cogn. Neurosci.* 25 (1), 22–28.
- Nickel, M., Gu, C., 2018. Regulation of central nervous system myelination in higher brain functions. *Neural Plast.* 2018, 6436453.
- Nyberg, L., et al., 2012. Memory aging and brain maintenance. *Trends Cogn. Sci.* 16 (5), 292–305.
- Ofen, N., et al., 2007. Development of the declarative memory system in the human brain. *Nat. Neurosci.* 10 (9), 1198–1205.
- Ostby, Y., et al., 2009. Heterogeneity in subcortical brain development: a structural magnetic resonance imaging study of brain maturation from 8 to 30 years. *J. Neurosci.* 29 (38), 11772–11782.
- Ostby, Y., et al., 2012. Dissociating memory processes in the developing brain: the role of hippocampal volume and cortical thickness in recall after minutes versus days. *Cereb. Cortex* 22 (2), 381–390.
- Pereira, J.B., et al., 2014. Regional vulnerability of hippocampal subfields to aging measured by structural and diffusion MRI. *Hippocampus* 24 (4), 403–414.
- Poppenk, J., Moscovitch, M., 2011. A hippocampal marker of recollection memory ability among healthy young adults: contributions of posterior and anterior segments. *Neuron* 72 (6), 931–937.
- Poppenk, J., et al., 2013. Long-axis specialization of the human hippocampus. *Trends Cogn. Sci.* 17 (5), 230–240.
- Reuter, M., et al., 2015. Head motion during MRI acquisition reduces gray matter volume and thickness estimates. *Neuroimage* 107, 107–115.
- Riggins, T., et al., 2015. Developmental differences in relations between episodic memory and hippocampal subregion volume during early childhood. *Child Dev.* 86 (6), 1710–1718.
- Riggins, T., et al., 2018. Protracted hippocampal development is associated with age-related improvements in memory during early childhood. *Neuroimage* 174, 127–137.
- Sala, G., Gobet, F., 2017. Experts' memory superiority for domain-specific random material generalizes across fields of expertise: a meta-analysis. *Mem. Cognit.* 45 (2), 183–193.
- Sastre 3rd, M., et al., 2016. Age- and performance-related differences in hippocampal contributions to episodic retrieval. *Dev. Cogn. Neurosci.* 19, 42–50.
- Schlichting, M.L., et al., 2017. Hippocampal structure predicts statistical learning and associative inference abilities during development. *J. Cogn. Neurosci.* 29 (1), 37–51.
- Scoville, W.B., Milner, B., 1957. Loss of recent memory after bilateral hippocampal lesions. *J. Neurol. Neurosurg. Psychiatry* 20 (1), 11–21.
- Selmecey, D., et al., 2019. Longitudinal trajectories of hippocampal and prefrontal contributions to episodic retrieval: effects of age and puberty. *Dev. Cogn. Neurosci.* 36, 100599.
- Shing, Y.L., et al., 2010. Episodic memory across the lifespan: the contributions of associative and strategic components. *Neurosci. Biobehav. Rev.* 34 (7), 1080–1091.
- Small, S.A., et al., 2011. A pathophysiological framework of hippocampal dysfunction in ageing and disease. *Nat. Rev. Neurosci.* 12 (10), 585–601.
- Smith, S.M., et al., 2004. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23 (Suppl. 1), S208–S219.
- Sneve, M.H., et al., 2015. Mechanisms underlying encoding of short-lived versus durable episodic memories. *J. Neurosci.* 35 (13), 5202–5212.
- Spaniol, J., et al., 2009. Event-related fMRI studies of episodic encoding and retrieval: meta-analyses using activation likelihood estimation. *Neuropsychologia* 47 (8–9), 1765–1779.
- Squire, L.R., Alvarez, P., 1995. Retrograde amnesia and memory consolidation: a neurobiological perspective. *Curr. Opin. Neurobiol.* 5 (2), 169–177.
- Strange, B.A., et al., 2014. Functional organization of the hippocampal longitudinal axis. *Nat. Rev. Neurosci.* 15 (10), 655–669.
- Tamnes, C.K., et al., 2013. Brain development and aging: overlapping and unique patterns of change. *Neuroimage* 68, 63–74.
- Tamnes, C.K., et al., 2014. Regional hippocampal volumes and development predict learning and memory. *Dev. Neurosci.* 36 (3–4), 161–174.
- Tatu, L., Vuillier, F., 2014. Structure and vascularization of the human hippocampus. *Front. Neurol. Neurosci.* 34, 18–25.
- van Norden, A.G., et al., 2012. Diffusion tensor imaging of the hippocampus and verbal memory performance: the RUN DMC study. *Hum. Brain Mapp.* 33 (3), 542–551.

- Van Petten, C., 2004. Relationship between hippocampal volume and memory ability in healthy individuals across the lifespan: review and meta-analysis. *Neuropsychologia* 42 (10), 1394–1413.
- Vidal-Pineiro, D., et al., 2018. Maintained frontal activity underlies high memory function over 8 years in aging. *Cereb. Cortex*.
- Walhovd, K.B., et al., 2004. Size does matter in the long run: hippocampal and cortical volume predict recall across weeks. *Neurology* 63 (7), 1193–1197.
- Wechsler, D., 1999. Wechsler Abbreviated Scale of Intelligence. The Psychological Corporation, San Antonio, TX.
- Wechsler, D., 2008. Wechsler Preschool and Primary Scale of Intelligence - III, Norwegian Version. Harcourt Assessment.
- Wolf, D., et al., 2015. Differential associations of age with volume and microstructure of hippocampal subfields in healthy older adults. *Hum. Brain Mapp.* 36 (10), 3819–3831.
- Wood, S.N., 2006. Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC.