Effects of Structural and Functional Prefrontal Cortex Maturation on Verbal Memory Development

Nadia Klijn[1]
Supervisors: Nils Müller[1,2], James McQueen[1,3,4], Guillén Fernández[1,2]

[1] Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands
[2] Radboud University Medical Centre Nijmegen, Donders Institute of Brain, Cognition and Behaviour, The Netherlands
[3]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[4] Behavioural Science Institute, Radboud University Nijmegen, The Netherlands

The prefrontal cortex, including core regions of our mnemonic system, is characterised by protracted structural and functional maturation from late childhood into early adulthood. We hypothesize that this relatively late maturation contributes to the steep increases in verbal memory performance observed during adolescence. To test this, we compared verbal memory scores of children (10-12 years old), adolescents (18 years old) and adults (25-32 years old). As a measure of structural maturation, we investigated whether cortical thickness of the left inferior frontal gyrus (LIFG, important for verbal memory encoding) mediates the effect of age on verbal memory performance. Furthermore, we investigated whether a developmental increase in specialisation (i.e., lateralisation) of functional connectivity between the inferior frontal gyrus (IFG) and the left medial temporal lobe mediates the effect of age on verbal memory performance. Firstly, our results show increased verbal memory performance and leftward lateralisation of functional connectivity, and decreased cortical thickness with age. More importantly, our results show that LIFG cortical thickness does, but LIFG functional specialisation does not significantly mediate the increase in verbal memory performance with age. This indicates that structural maturation is indeed an important contributor to verbal memory development during adolescence, whereas functional LIFG specialisation seems to play a less significant role.

Keywords: prefrontal cortex, maturation, adolescence, verbal memory, cortical thickness, functional connectivity

---

Corresponding author: Nadia Klijn; **E-mail:** nadiaklijn@hotmail.com

An intriguing question in the cognitive neuroscience of memory is how the maturation of neural structure and function contributes to the development of an adult-like declarative memory system. One central characteristic of declarative memory is our capacity to form new semantic representations, for instance by acquiring new words. But why exactly does our ability to remember words improve so drastically when we grow up, while our brain becomes less plastic at the same time? Brain regions supporting memory encoding undergo substantial functional and anatomical changes throughout the lifespan (e.g., Johnson, 2001; Simons & Spiers, 2003). Therefore, we expect that structural and functional differences between the developing and the adult brain underlie the differences in memory performance between these two. We investigated whether this is the case by comparing verbal memory scores of children (10-12 years old), adolescents (18 years old) and adults (25-32 years old). We examined whether these scores can be predicted by structural thinning and functional specialisation of brain areas that are related to verbal memory performance in adults.

Verbal memory: brain and behaviour

Our capacity to successfully remember information improves drastically across development with declarative memory only reaching full functionality in early adulthood (Baltes, Lindenberger & Staudinger, 2006; Finn et al., 2016). Within the declarative memory system, there is a transition from favouring perceptual, episodic-like representations towards favouring more abstract, semantic representations (Ofen & Shing, 2013). In line with this, verbal learning studies have revealed an increase in verbal memory performance across development (e.g., Vakil & Blachstein, 1997; Vakil, Blachstein & Sheinman, 1998; Blachstein & Vakil, 2016; Davis et al., 2013). Using the Rey Auditory Verbal Learning Test, these studies measured immediate and delayed recall after intentional encoding trials. Results of these studies indicate an increase in verbal memory performance during childhood and adolescence that peaks in early adulthood, after which performance levels decrease again with age. However, in these studies participants were aware that their memory for the items would be tested later on in the experiment. Therefore, differences in memory performance across age groups are not only driven by differences in mnemonic abilities, but are also influenced by differences in strategic abilities between children and adults (Shing, Werkle-Bergner, Li & Lindenberger, 2008). To truly assess the pure mnemonic advances in verbal memory development, it is important to disentangle the (neural) maturational changes underlying true mnemonic abilities from these strategic abilities. To avoid the confound of unequal strategy use in the current study, we investigated the developmental differences in performance on a verbal memory task in which encoding is incidental and thus strategy-free.

The maturational trajectories of neural regions involved in verbal memory performance in adults can provide us with insights on changes in structural and functional properties that potentially underlie differences in memory performance across development. Studies investigating verbal memory in the adult brain found that subsequent memory for verbal material is primarily associated with left lateralized activation in the medial temporal lobe (MTL) and the prefrontal cortex (PFC), more specifically the left inferior frontal gyrus (LIFG) (e.g., Wagner et al., 1998; Golby et al., 2001; Opitz, Mecklinger & Friederici., 2000; Duverne, Motamedinia & Rugg, 2009; see for a meta-analysis Kim, 2011). Activation in the anterior extent of the LIFG (Brodmann's area [BA] 45/47) is assumed to reflect semantic control processes mediating internal representations. These processes then optimize memory encoding when interacting with core storage regions in the MTL (Thompson-Schill, D'Esposito, Aguirre & Farah, 1997; Bokde, Tagaments, Friedman & Horwitz, 2001; Simons & Spiers, 2003; Kim, 2011). Furthermore, activation of MTL regions is assumed to be involved in binding together the features of an event into a compound representation and in consolidating this representation for later use (Werkle-Bergner, Müller, Li & Lindenberger, 2006). Structural maturation of these regions and the maturation of their interaction with each other, could therefore potentially explain differences in verbal memory performance across development. Nevertheless, it is good to keep in mind here that the correlational evidence mentioned above only provides us with limited information on the cause of the reported effects.

Structural maturation – Cortical thickness

Trajectories of structural PFC maturation differ substantially from that of other cortical regions, such as the MTL (Sowell et al., 2003; Lenroot & Giedd, 2006; Raz & Rodrigue, 2006). Anatomical evidence suggests that most MTL regions mature at a relatively fast rate before late childhood, whereas the PFC and associated neural networks undergo serious age-related changes well into adolescence and early adulthood (e.g., Sowell et al., 2003; Sowell et al., 2004; Gogtay et al., 2004; Gogtay et al., 2006). Because of this differential pattern of maturation between the MTL and the PFC, the increase in verbal memory performance from late childhood into early adulthood as suggested by the behavioural literature is likely driven by ongoing maturational processes in the PFC. One protracted structural change in the PFC that could be of relevance in explaining developmental differences in verbal memory performance, is the gradual decline in cortical thickness in frontal association cortices until early adulthood (Sowell et al., 2003;Sowell et al., 2004; Gogtay et al., 2004; Koolschijn & Crone, 2013; Wierenga, Langen, Oranje & Durston, 2014; Amlien et al., 2014; Ducharme et al., 2016). Decreased cortical thickness over time is usually associated with increased synaptic pruning from childhood into adulthood (Petanjek et al., 2011). Because of this, cortical thinning can be indirectly associated with more efficient processing, which makes it a relevant measure in relation to gains in memory performance during development (Østby, Tamnes, Fjell & Walhovd, 2011 ; Sowell, Delis, Stiles & Jernigan, 2001). Sowell et al. (2001) for instance showed that an increase in verbal memory performance across development is associated with decreasing cortical thickness of the PFC, but less so of the MTL. However, Sowell et al. (2001) used an intentional verbal learning task to test their hypotheses. Therefore, based on this study we cannot conclude whether PFC structural maturation is specifically beneficial for verbal memory development itself, or whether the effect is driven only by improved strategic abilities that also rely on the PFC. Since our task is strategy-free, we would be able to draw such conclusions when testing the hypothesis that protracted cortical thinning of the PFC allows for increasing verbal memory performance.

Functional maturation – Specialisation of connectivity

On a functional level, protracted synaptic pruning of the PFC from childhood into early adulthood (Petanjek et al., 2011) potentially causes a decrease in short-range, inter-regional, functional connectivity (segregation) while simultaneously causing long-range connectivity to keep increasing due to task-related co-activation, known as integration. (Jolles, Van Buchem, Crone & Rombouts, 2011). This is in line with the finding that children show more diffuse functional connectivity patterns and increased short-range functional connectivity compared to adults, whereas adults show increased long-range functional connectivity patterns (Kelly et al., 2009). Therefore, protracted synaptic pruning of the PFC, as structurally reflected in protracted cortical thinning across development, is potentially also associated with specialisation of brain regions within larger functional networks (Durston & Casey, 2006; Durston et al., 2006; Fair et al., 2007;Fair et al., 2009; Supekar, Musen & Menon, 2009; reviewed in Johnson, 2011). In studies using language-related tasks, this functional specialisation of brain regions is often observed as decreases in bilaterality of task-related fMRI activity towards a more left lateralized activation pattern from childhood to early adulthood. This lateralisation is then associated with an increase in performance on verbal tasks (Holland et al., 2001; Szaflarski, Holland, Schithorst & Byars, 2006a; Szwaflarski et al., 2006b: Ressel et al., 2008; Everts et al., 2009; reviewed in Holland et al., 2007). Interestingly, Szaflarski et al. (2006a) show an inverted u-shape for language lateralisation similar as to what has been shown in behavioural studies investigating verbal memory across the lifespan (Blachstein & Vakil, 2016; Davis et al., 2013), where the learning curves increase until age 20, plateaus around age 25 and slowly decreases afterwards. However, most of these studies used a verb generation task which does not allow us to draw conclusions on whether maturing lateralisation of language functions predicts verbal memory development. Contrarily to previous language-related studies exploring developmental lateralisation of task-related brain activity, we examined lateralisation of resting-state functional connectivity between IFG – left MTL regions with age. This enabled us to investigate whether increasing functional specialisation (i.e., lateralisation) of the PFC between childhood and early adulthood is related to increasing verbal memory capacities without the fMRI measurements being confounded by differences in task performance between groups. Clearly, these measures do not allow us to make direct inferences about specialisation in response to the task. Resting-state functional connectivity is however believed to reflect recent experience as well as neuroanatomy and thus provides insight on both levels (Uddin, Supekar & Menon, 2010; Tavor et al., 2016). This makes it an appropriate measure to detect whether maturing functional specialisation of core mnemonic regions contributes to the development of an adult-like memory system.

Current study

In this study we are taking the behavioural, structural and functional maturational patterns as described above together. To do so, we carried out an implicit verbal memory task, and both structural and functional (resting-state) MRI scans in children (10-12 years old), adolescents (18 years old) and adults (25-32 years old). This selection of groups is most optimal since it provides us with participants in the beginning, in the middle and at the end of the protracted cortical thinning process of the PFC (Sowell et al., 2003; Sowell et al., 2004; Gogtay et al., 2004). With this data, our main goal was to investigate whether protracted structural and functional maturation of the PFC from late childhood into early adulthood can (at least partially) explain the steep increase in verbal memory performance across these developmental stages.

Our first research question involves the behavioural level: can we replicate the effect of age group on verbal memory scores?

The second question we aim to answer is whether cortical thickness of the left anterior IFG (BA 45/47, further referred to as LIFG) indeed decreases decreases as a function of age between the groups, as an indirect measure of the amount of synaptic pruning that has taken place allowing for greater processing efficiency. We also investigated whether this decrease in cortical thickness mediates the hypothesised increase in verbal memory performance with age.

Our third and last question is whether the specialisation (i.e., lateralisation) of associated LIFG functional connectivity indeed increases with age group, and whether this increase mediates the hypothesised increase in verbal memory performance with age. Lateralisation of LIFG involvement was measured as its resting-state functional connectivity with the already matured left MTL relative to the resting-state functional connectivity of the RIFG with the left MTL.

Materials and Methods

Participants

Ninety right-handed native Dutch-speaking volunteers participated in this study. Thirty of them were adults aged between 25-32 years old (M_{age} = 26.9 years, *SD* = 21.9 months, 12 male), 29 were adolescents aged 18 (M_{age} = 18.5 years, *SD* = 3.1 months, 10 male) and 31 were children aged between 10-12 years old (M_{age} = 11.0 years *SD* = 8.8 months, 8 male). All subjects had normal hearing and normal or corrected-to-normal vision. Furthermore, participants were required to have no history of injury or disease known to affect the central nervous system function (including neuropsychological disorders such as dyslexia, autism and ADHD) and to not have MRI contraindications. Adults and adolescents were recruited from the student population of Radboud University, Nijmegen, and from the surrounding community. Children were recruited through presentations and flyers at local schools. The study was approved by the local ethics committee, the CMO Arnhem – Nijmegen. Written informed consent was obtained prior to participation from all participants who were at least age 18 and from both parents of participants under age 18.

Of all 90 participants, 89 participants ultimately satisfied the inclusion criteria for the behavioural analysis (one child had to be excluded due to dyslexia). Of these, 84 had adequate processed and quality-checked MRI data for the cortical thickness analyses (30 adults, 26 adolescents and 28 children) and 84 had adequate data for the functional connectivity analyses (28 adults, 28 adolescents and 28 children).

Experimental stimuli

The experimental stimuli comprised recordings of 72 bi- and trisyllabic Dutch verbs geared to the children's level of interaction with the world. These words were selected based on a short pilot in which seven children aged between 10-12 years old got a list of 78 hand and arm action verbs. These words were selected from a list of Dutch verbs based on their meaning (a relatively basic action carried out by our hands or arms, like 'throwing') and on the amount of syllables (two or three). On this list of 78 verbs, the children had to circle the words for which they did not know the meaning. If an action verb was encircled twice, we did not include it in our stimulus list. All verbs included in our study were therefore assumed to be well known for both our adult participants as well as for the children.

These stimuli were recorded by a male and a female speaker to introduce variation and create a more natural linguistic encoding situation during the experiment. We chose an auditory stimulus presentation to prevent an encoding advantage towards the older age groups due to better reading abilities. For each stimulus, three different versions with different intensities were created to enable a loudness rating task in the encoding phase. This resulted in one version of 65 decibel (dB), one version of 67.5 dB and one version of 70 dB for each word. We counterbalanced the presentation of one of these versions for a specific word across subjects. Furthermore, word frequency, voice and the amount of syllables were balanced across target ('old', to be remembered words) and filler ('new', words added in the recognition phase) conditions. The distribution of the individual words over these two conditions was randomized across subjects. In the encoding phase 36 verbs were presented. In the recognition phase the same 36 verbs as well as 36 new, unstudied, verbs were presented to the participants. For the lists of the recognition phase, words were roughly equally spaced relative to their first occurrence in the encoding phase, and the sequences of words were pseudorandomly ordered such that no more than three items belonging to the same condition (old versus new) occurred sequentially.

Verbal memory task

The verbal memory task consisted of two separate blocks run in Presentation (Version 16, www.neurobs.com). The first block consisted of an incidental, shallow encoding task. We chose a shallow encoding phase (loudness rating task) to ensure that children did not have a disadvantage at the level of task performance during the study block that might be present in a more difficult task, such as a deep encoding semantic decision task. More difficulty with the encoding task could namely confound children's mnemonic abilities. Furthermore, we chose incidental encoding to rule out the possibility that differences in memory performance across groups were mainly attributable to differences in strategic encoding abilities (Shing et al., 2010). The encoding phase was immediately followed by a second block in which recognition memory of the verbs was tested.

Before the start of the first block, participants were instructed that words would be presented to them via headphones and that their task was to rate the loudness of the presented stimulus in comparison to the loudness of the previous stimulus (the system's sound volumes were kept constant across participants). To rate the loudness of the verbs, participants could use six different responses presented on a computer screen also reflecting their degree of confidence, namely: 'surely louder', 'louder', 'maybe louder', 'maybe softer', 'softer', 'surely softer'. Whether the options for 'louder' and 'softer' were presented on the left or on the right was counterbalanced across subjects. The participants were instructed to use their left and right index, middle and ring fingers and let them rest on the response buttons on the keyboard (buttons 1-6 in the upper-left corner). Instructions emphasized the need for both speed and accuracy. They were not informed that their memory for the verbs would be tested afterwards. The incidental encoding phase consisted of 36 unique trials in total, such that each target verb was encountered only once before test.

Immediately after completing this encoding phase, the participants got the instructions for the recognition task. They were told that they would hear the words again together with words they had not heard in the first block of the experiment. All words in the recognition phase were presented at the 67.5 dB level. We explained to the participants that their task was to judge which words they already encountered ('old') and which words they did not ('new'), and to indicate their degree of confidence on this decision in the same way as during the encoding phase. The assignment of old and new responses to the left or right hand was counterbalanced across participants. Furthermore, the participants were instructed to respond as in the first block, and again as fast and as accurate as possible. The recognition phase consisted of 72 trials in total, of which 36 were old verbs and 36 were new (counterbalanced across subjects).

Both in the encoding and the recognition phase, each trial comprised of a white fixation cross in the center of the screen presented for 200 ms, followed by a visual presentation of the response continua for 300 ms before presentation of the stimulus. After the onset of the word, participants could press a response button. If this response was given within 1500 ms after stimulus onset, the participant had to wait until 1500 ms after stimulus onset had passed before the fixation cross was presented again to make sure participants could not go through the whole experiment without listening to the words. The inter-trial interval during which the fixation cross was presented was 500 ms.

Procedure

The verbal memory task reported here was implemented within a bigger study investigating the role of prior knowledge in schema memory development (Müller et al., in preparation). For this study, a battery of tests was performed. On Day 1, participants came to the Donders Centre for Cognitive Neuroimaging and started in the behavioural lab with a practice fractal N-back task (Ragland et al., 2002), after which the described verbal memory task was performed. Immediately after completing the verbal memory task, participants were taken to the MRI lab where we first made their resting-state fMRI scan. Then the fractal N-back task that was practiced before was performed in the scanner and a structural scan was conducted. Afterwards, participants performed five sessions of a memory game in the lab and at home, after which they came back to the Donders Centre for Cognitive Neuroimaging on Day 8. Then they performed the recall session of the memory game in the MRI scanner after which another structural scan was made. For our analyses, we used either the structural scan from Day 1, the structural scan from Day 8, or an average of the two, based on which of these three options allowed for the best quality (i.e., least motion artifacts). After the second MRI session, the procedure continued with a Wisconsin Card Sorting Task (Heaton, Chelune, Talley, Kay & Curtiss, 1993), a Digit Span task (Alloway, 2007) and a long-delay follow-up on the memory game for adolescents and adults.

fMRI data acquisition

Scanning was performed with a Siemens Magnetom Skyra 3 tesla MR scanner equipped with a 32-channel phased array head coil. First, the resting-state scans were acquired during which the participants were instructed to keep their eyes fixated on a black fixation cross in the center of a white screen. Furthermore, we told them to try to think of nothing in particular. A total of 900 blood oxygen level-dependent scans were acquired using a T_2*-weighted gradient-echo, whole-brain echo planar imaging (EPIs) sequence with the following parameters: time repetition (TR) = 657 ms; time echo (TE) = 36.80 ms; flip angle = 50°; multiband factor = 8; matrix size = 88 x 88; field-of-view (FOV) = 210 mm; slice thickness = 2.4 mm; no slice gap; 64 slices acquired interleaved. The first ten scans were discarded to allow for equilibration of T_1 saturation effects. Furthermore, to control for the effect of field inhomogeneities, we acquired a fieldmap with the same parameters as the resting-state scans (TE1 = 4.54 ms, TE2 = 7.00 ms).

After acquiring the functional data, a high resolution T1-weighted anatomical scan was made to enable registration and the analysis of cortical thickness. We used a magnetization prepared, rapid-acquisition gradient echo sequence (parameters: TR = 2300 ms; TE = 3.03 ms; flip angle = 8°; matrix size = 256 x 256; FOV = 256 mm; slice thickness = 1 mm; 192 sagittal slices).

Data analysis

Behavioural data analysis – Verbal memory. On a behavioural level, we wanted to investigate whether we could replicate the effect of age group on verbal memory scores as described in the literature.

Statistical tests were performed using IBM SPSS Statistics (23.0, SPSS Inc., Chicago, USA). As a first control analysis, we tested whether the participants could generally discriminate successfully between old and new verbs. For this, the recognition memory data was first analysed using a 2 x 6 ANOVA with the 'Condition' of the items as the first factor (old versus new), the confidence rating as the other six-level factor ('Response') and the proportion of responses given for this combination of factors as the dependent variable (Liu, Qin, Rijpkema, Luo & Fernandez, 2010). We then applied post-hoc paired samples t-tests to verify for each response bin whether participants could discriminate old from new items. If the difference between the proportions of old and new responses for a certain confidence bin was proven significant, we concluded that most of the responses in this bin reflected true memory and decided to include responses from that bin for further analyses.

To achieve a representative score for memory performance per participant, we calculated the z-scores of both hit (old items recognized as old) and false alarm rates (new items recognized as old). Then, the z-scores for false alarms were subtracted from the z-scores for the hits (MacMillan & Creelman, 2005), leading to a d', representing the memory performance for each subject while controlling for response biases. This d' served as the dependent variable in a one-way ANOVA and pairwise comparison with age group as an independent variable, to test whether there is a significant difference between age groups on performance in a verbal memory task. To add meaning to the development of d' memory scores, we also performed one-way ANOVA's and pairwise comparison to investigate whether and how hit rates and false alarm rates differ across age groups.

Structural MRI data analysis – Cortical thickness. Analysis of the structural MRI data served two purposes. Firstly, we wanted to determine whether cortical thickness of the LIFG indeed decreases with age group (as an indirect measure of the amount of synaptic pruning that has taken place, allowing for greater processing efficiency). Furthermore, we hypothesized that this decrease in cortical thickness mediates the increase in verbal memory performance with age group.

Preprocessing. High resolution T_1-weighted anatomical scans were analysed to estimate cortical thickness using FreeSurfer 5.1 software (http://surfer.nmr.mgh.harvard.edu/). This procedure provided us with a measure of cortical thickness for each participant at each point of the reconstructed surface allowing for the detection of submillimeter differences between age groups (e.g., Fjell et al., 2010; Fischl & Dale, 2000). The automated reconstruction involved motion correction, removal of non-brain tissue (Clarkson et al., 2011), intensity normalization (Sled, Zijdenbos & Evans, 1998), tessellation of gray/white matter boundary automated topology correction (Segonne, Pacheco & Fischl, 2007), surface deformation, and registration, segmentation of subcortical white matter and deep brain structures (Hutton, Draganski, Ashburner & Weiskopf, 2009). Estimates of cortical thickness on the tessellated surface were obtained by calculating the closest distance between representations of the cortical surface and the gray/white matter border (Fischl & Dale, 2000). After construction of surface-based maps for each participant's anatomical scan, cortical structures of each individual scan were registered to a spherical atlas and labeled using the Desikan-Killiany atlas (Desikan et al., 2006). This resulted in estimates of average cortical thickness in millimeters for every labeled cortical structure per participant.

Analyses. Cortical thickness analyses were performed, extracting each participant's cortical thickness of the anterior extent of the LIFG (BA45/47). This anterior extent including both the pars triangularis and the pars orbitalis which have been shown to contribute to verbal memory performance and to play a key role in semantic control (Bokde et al., 2001). Therefore, we extracted both the pars triangularis and the pars orbitalis of the LIFG (BA45/47). These regions of the anterior extent had to be extracted separately due to the arrangement of the Desikan-Killiany atlas used by FreeSurfer 5.1. Based on the described literature we were interested in the anterior extent of the LIFG as a whole, we collapsed these Regions of interest (ROI) resulting in a mean of the average cortical thickness estimates of both regions. The result of this is our measure of average anterior LIFG cortical thickness for each participant.

An ANOVA and pairwise comparison were performed to test the effects of age group on cortical thickness. We also performed a simple linear regression to test whether there was an overall effect of cortical thickness on d' memory scores (uncorrected for age group). Finally, a mediation analysis with the cortical thickness of the LIFG and age group as regressors and verbal memory score (d') as a dependent variable was performed using PROCESS for SPSS (Hayes, 2013). This was to investigate whether the effect of age group on d' memory scores would be mediated by hypothesized decreasing cortical thickness of the LIFG with age.

Resting-state fMRI data analysis – Functional specialisation. The resting-state fMRI data were analysed to determine whether the specialisation of LIFG functional connectivity indeed increases with age group, and whether this increase also mediates the hypothesised increase in verbal memory performance with age.

Preprocessing. Preprocessing steps and analyses were performed in FSL (Jenkinson, Beckmann, Behrens, Woolrich & Smith, 2012) and consisted of several steps using a combination of fMRI Expert Analysis Tool Version 6.00 and custom MATLAB (8.6, The MathWorks Inc., Massachusetts, USA) scripts. Registration to high resolution structural space images was carried out using FLIRT (Jenkinson, Bannister, Brady & Smith, 2002). Registration from high resolution structural to standard space was then further refined using FNIRT nonlinear registration (Andersson, Jenkinson & Smith, 2007a; Andersson, Jenkinson & Smith, 2007b). We applied motion correction using MCFLIRT (Jenkinson et al., 2002), nonbrain removal using BET (Smith, 2002), spatial smoothing using a Gaussian kernel of 5 mm and grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor. To correct for motion, we then used an Independent Component Analysis-based Automatic Removal of Motion Artifacts (ICA-AROMA) allowing us to preserve the autocorrelation structure of fMRI time-series and to avoid heteroscedasticity in group-level statistics (Pruim et al., 2015). ICA-AROMA uses an ICA decomposition on the partly preprocessed fMRI data. In this decomposition, a set of spatial and temporal features and a classification procedure were used to identify independent components representing motion artifacts. Finally, the selected components were removed from the fMRI time-series through linear regression. Residual noise was then removed by the regression of white matter (WM) and cerebrospinal fluid (CSF). The mean WM and CSF signals were extracted using masks obtained by multiplying a participant-specific tissue prior with an MNI152-derived tissue prior, both thresholded at 95% tissue probability (Pruim et al., 2015). Hereafter, we used high-pass temporal filtering to remove slow drifts (sigma = 100 s). Resting-state scans were co-registered to the participant's structural T_1-weighted image.

Analyses. Functional connectivity analyses were performed measuring the correlations between the following ROIs as explained earlier: 1) between the LIFG (-50,25,12 in MNI space, BA 45/47, see Wagner et al., 1998) and the left MTL (737 voxels, center of gravity = (74,46,35)) and 2) between the right anterior IFG (from now on RIFG, 50,25,12) and the left MTL. For the left and right anterior IFG, we created an 8 mm sphere around the MNI coordinates from Wagner et al. (1998) and their right counterparts that served as a mask to warp onto the functional resting-state data of each subject. As a target in the left MTL, we used the center of an MTL cluster which exhibited a strong resting-state connectivity to the LIFG ($r \geq 0.2$) in an analysis of 1000 subjects (Yeo et al., 2011; via www.neurosynth.org).

We extracted the first eigenvariates for each ROI in native space by warping the three masks from MNI space into functional space. As a measure of functional connectivity, we calculated the Pearson correlations between the resting-state time series of the LIFG and the left MTL and between the RIFG and the left MTL for each participant. This resulted in two connectivity values per participant that we used to compute a laterality index (LI) for each participant. This LI was defined as the difference between the correlational value of the LIFG with the left MTL and the correlational value of the RIFG with the left MTL (r left – r right) (Nielsen, Zielinkski, Ferguson, Lainhart & Anderson, 2013). An ANOVA and pairwise comparison were performed to test the effects of age group on LI. We also performed a simple linear regression to test whether there was an overall effect of LI on d' memory scores (uncorrected for age group). If such an effect would exist, we would perform a mediation analysis with LI and age group as regressors and verbal memory score (d') as a dependent variable using PROCESS for SPSS (Hayes, 2013). In this way, we wanted to investigate whether the effect of age group on d' memory scores is mediated by possibly increasing lateralisation of functional connectivity with age. Additionally, a complementary simple linear regression was performed to investigate whether the assumed relationship between cortical thinning and functional specialisation as assumed in the literature indeed exists (Johnson, 2011).

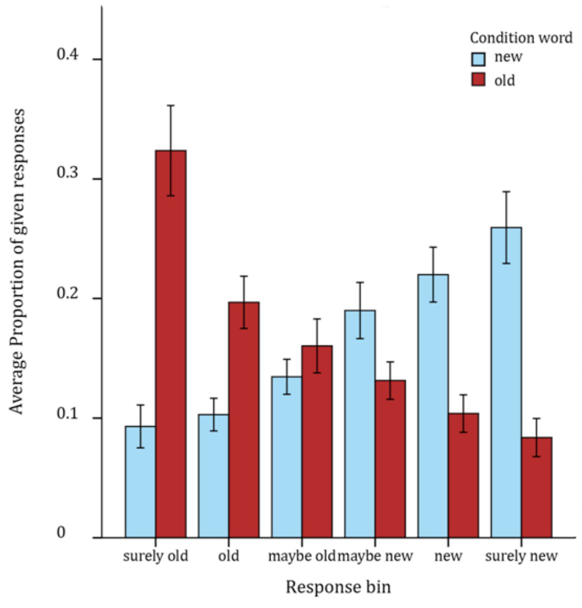

Fig. 1. Mean proportion of responses given for 'old' and 'new' conditions in each response bin.

Results

Behavioural results – Verbal memory

An interaction effect between the factors 'Condition' (old or new) and 'Response' on the proportion of responses given (2 x 6 ANOVA) showed that participants were overall well able to distinguish between old and new verbs in the recognition phase of the verbal memory test ($F(5, 1068) = 89.41$, $p < .001$). See Table 1 and Figure 1 for the mean proportions of responses given in 'old' and 'new' conditions for each response bin.

Post-hoc pairwise comparisons corrected for multiple comparisons (Bonferroni) revealed that the proportion of trials with responses 'surely old' and 'old' was higher for old verbs than for new verbs (all $p < .001$). On the contrary, responding with 'maybe new', 'new' and 'surely new' occurred more often during trials with new verbs than with old verbs (all $p < .001$). A non-significant difference between proportions of old and new trials receiving the response 'maybe old' ($p > .05$) indicates that participants could not discriminate significantly between old and new verbs when they used this response button. Because of this, verbs that were responded to with 'maybe old' were categorized as trials of no interest and were excluded from further analyses consistent with previous studies (Liu et al., 2010; Qin et al., 2009).

Mean false alarm rates, hit rates and d' memory

Table 1.

Mean proportion of responses given for 'old' and 'new' conditions in each response bin, with standard deviations.

Response button	Condition 'new'	Condition 'old'
Sure old	0.093(0,085)	0.324(0,180)
Old	0.103(0,065)	0.197(0,104)
Maybe old	0.135(0,070)	0.160(0,108)
Maybe new	0.190(0,112)	0.131(0,075)
New	0.220(0,109)	0.104(0,075)
Sure new	0.259(0,143)	0.084(0,076)

scores for the different age groups are shown in Table 2. Firstly, we wanted to investigate whether there is a main effect of age group on mean d' scores using a one-way ANOVA. We found that age group indeed affects mean d' scores (*F*(2, 86) = 5.15, *p* < .01, see Table 2). Pairwise comparisons showed that mean d' scores for children were significantly lower than those for adolescents and adults (*p* = .003 and *p* = .017, respectively), whereas adolescents' mean scores were not significantly higher than those for adults (*p* > .05, see also Fig. 2A). These effects seem to be driven mostly by differences in mean false alarm rates between age groups (*F*(2, 86) = 8.20, *p* = .001), whereas mean hit rates did not differ significantly between children, adolescents and adults (*F*(2, 86) = 0.10, *p* = 091). Pairwise comparisons confirmed that mean false alarm rates follow an inverted pattern of the d' scores where children's false alarm rates were significantly higher than those for adolescents and adults (*p* ≤ .01), whereas adolescents' mean scores were lower than those for adults (*p* > .05), but again this difference was not significant.

Structural MRI results – Cortical thickness

In a mediation analysis we wanted to investigate the research question whether cortical thickness of the LIFG predicts memory scores across age groups, as an indirect measure of the amount of synaptic pruning that has taken place allowing for greater processing efficiency. We already showed that age group predicts memory scores. Then, to check whether the decrease in cortical thickness in the LIFG from childhood into adulthood as reported in previous studies (e.g., Amlien et al., 2014) was also applicable to our sample, we performed a one-way ANOVA revealing an effect of age group on the absolute cortical thickness of the LIFG (*F*(2, 81) = 18.67, *p* < .001, see Table 2). This effect reflected a decrease in cortical thickness with increasing age group (see Fig. 2B). Pairwise comparisons showed that cortical thickness of the LIFG was significantly higher for children than for adolescents and adults (*p* < .005 and *p* < .001, respectively). There was also a significant decrease from adolescence into adulthood (*p* < .01).

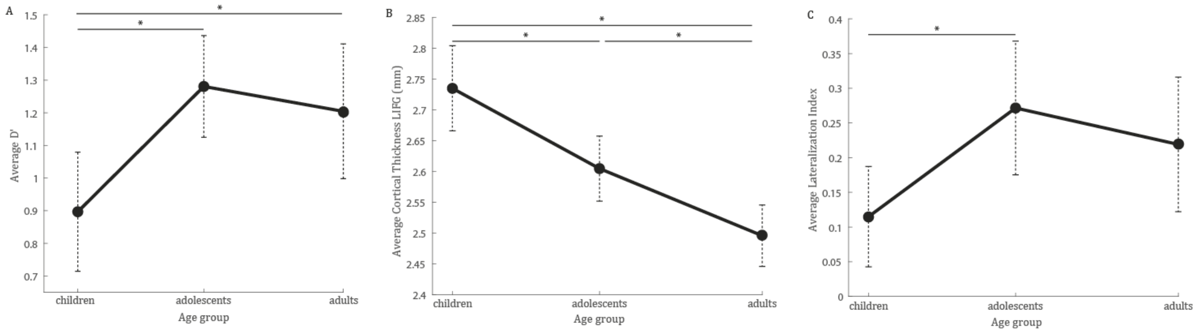

Fig. 2A. Mean d′ per age group. Children's scores are significantly lower than that of adolescents and adults. **B.** Mean cortical thickness of the LIFG per age group. Cortical thickness significantly decreases with age across all age groups. **C.** Mean laterality indices of IFG-left MTL connectivity per age group. Children's IFG – left MTL connectivity is significantly less left-lateralized than adolescents.

Table 2.

Means and standard deviations for each variable per age group.

	Children	Adolescents	Adults
Hit rate	0.608(0.123)	0.617(0.145)	0.624(0.139)
False alarm rate	0.288(0.123)	0.180(0.082)	0.213(0.108)
d' memory scores	0.897(0.488)	1.281(0.409)	1.204(0.554)
Cortical thickness LIFG (mm)	2.735(0.178)	2.605(0.131)	2.496(0.133)
Laterality index	0.115(0.186)	0.272(0.249)	0.219(0.250)
LIFG – left MTL connectivity	0.178(0.179)	0.377(0.203)	0.282(0.235)

A simple linear regression showed that cortical thickness of the LIFG predicts d' memory scores irrespective of age group ($F(1, 82) = 9.25$, $p < .005$, $R^2 = .10$, $\beta = -.93$).

Because we found main effects of age group on d' and cortical thickness and of cortical thickness on d' (uncorrected for group), we performed a mediation analysis to investigate whether increasing age group leads to a better verbal memory performance because of the mediating effect of maturing LIFG cortical thickness. This showed that age group was no longer a significant predictor of d' scores after controlling for the mediating factor cortical thickness of the LIFG (age group: $\beta = .07$, $SE = 0.08$, $p > .05$; Cortical thickness: $\beta = -.74$, $SE = 0.37$, $p < .05$). The indirect effect of age group on d' scores through LIFG cortical thickness was tested using a bootstrap estimation approach with 1000 samples using PROCESS (Shrout & Bolger, 2002; Hayes, 2013). The results indicated the indirect coefficient was significant ($\beta = .09$, $SE = 0.05$, 95% confidence interval (CI) = 0.003, 0.21). This is an indication that cortical thickness of the LIFG mediates the effect of age group on d' scores. Approximately 11% of the variation was accounted for by age group and cortical thickness of the LIFG ($R^2 = 0.11$).

Resting-state fMRI results – Functional specialisation

In another mediation analysis we wanted to investigate the research question whether the degree of specialisation (i.e., lateralisation) of functional LIFG involvement predicts memory scores across age groups. To check whether the lateralisation of IFG functional connectivity actually increases from childhood into adulthood, we performed a one-way ANOVA revealing an significant effect of age group on the lateralisation index of functional connectivity of the IFG ($F(2, 81) = 3.36$, $p < .05$, see Table 2). Pairwise comparison showed that there was a significant increase in leftward lateralisation of IFG functional connectivity from childhood into adolescence ($p < .05$, see Fig. 2C). Adults also showed higher lateralisation indices than children, but this difference was not significant ($p > .05$), as was the decrease in lateralisation indices between adolescents and adults ($p > .05$).

Furthermore, a mediation analysis testing our hypothesis that maturation of functional specialisation (as reflected in an increasing lateralisation of IFG-left MTL connectivity) underlies the increase in verbal memory performance with age also requires an effect of laterality index on d'. However, a simple linear regression that was calculated to predict d' memory scores based on lateralisation of IFG-left MTL connectivity did not reach significance ($F(1, 82) = 0.41$, $p > .05$, $R^2 = 0.01$, $\beta = .16$). This means that we did not find evidence that lateralisation of IFG functional connectivity mediates the effect of age group on memory scores. Since cortical thickness did mediate verbal memory development, we performed a complementary simple linear regression to check whether the hypothesized relationship between decreasing cortical thickness of the LIFG and increasing specialisation (i.e., lateralisation) of IFG functional connectivity actually exists. There was a trend towards a significant effect from LIFG cortical thickness on laterality indices of IFG functional connectivity ($F(1, 79) = 3.82$, $p = .054$, $R^2 = 0.05$, $\beta = -.28$).

As an exploratory analysis, we further investigated whether LIFG connectivity with the left MTL in general could possibly be a predictor for developmental differences in verbal memory scores. There was an effect of age group on LIFG – left MTL connectivity ($F(2, 81) = 6.48$, $p < .01$, see Table 2), showing the same pattern as the lateralisation indices. This means there was an increase in LIFG

– left MTL functional connectivity from childhood into adolescence ($p < .05$). Adults also showed higher connectivity values than children, but this difference was not significant ($p > .05$), as was the decrease in connectivity between adolescents and adults ($p > .05$). Again, a simple linear regression showed no significant effect of LIFG – left MTL connectivity on d' memory scores ($F(1, 82) = 0.74$, $p > .05$, $R^2 = 0.10$, $\beta = .22$).

Discussion

Verbal memory

In this study, we aimed to investigate whether the maturation of neural structure and function is associated with behaviourally enhanced verbal memory performance from childhood into adulthood. Previous studies in the domain of verbal memory development that used intentional encoding tasks and a free recall test phase already reported that verbal memory performance increases during childhood and adolescence and flattens off in early adulthood, after which it decreases in older ages (e.g., Vakil & Blachstein, 1997; Vakil, Blachstein & Sheinman, 1998; Blachstein & Vakil, 2016; Davis et al., 2013). The results of our incidental verbal recognition memory task, reflecting strategy-free mnemonic abilities, are in line with these findings. Children between 10-12 years of age performed worse than 18 year old adolescents and 25-32 year old adults. Adults however did not differ significantly from adolescents, which is consistent with the flattening of the curve around these ages as shown in previous studies. Additionally, we showed that the increase in memory performance from childhood into adulthood is most probably driven by a decrease in false alarm rates between these age groups. This is consistent with early developmental word recognition studies showing a decrease in false memories throughout childhood and adolescence (e.g., Brainerd & Reyna, 1996; Brainerd, Reyna & Kneer, 1995; but see Brainerd, Reyna & Ceci, 2008 for an extensive discussion on this topic).

Cortical thickness

In examining the neural changes underlying memory development, we first investigated whether a developmental decrease in cortical thickness of the left inferior frontal gyrus (LIFG) mediates the effect of age group on verbal memory performance. We hypothesised that the protracted thinning of the PFC into early adulthood (e.g., Amlien et al., 2016), as an indirect measure of the amount of synaptic pruning that has taken place (Petanjek et al., 2011), would allow for greater processing efficiency in the LIFG over development. Verbal memory performance would benefit from this increased efficiency because the LIFG is associated with semantic control and has shown to be of particular importance in verbal memory tasks (Bokde et al., 2001; Kim, 2011). Our findings confirm this hypothesis.

Firstly, there was an effect of age group on average cortical thickness of the LIFG (BA 45/57). Children's LIFG were thicker than that of adolescents and adults, and LIFG cortical thickness kept decreasing from adolescence into adulthood. This linear pattern is consistent with several recent studies examining the developmental trajectories of cortical thickness in different brain areas in longitudinal and continuous cross-sectional samples (e.g., Koolschijn & Crone, 2013; Wierenga et al., 2014; Amlien et al., 2014; Ducharme et al., 2016). Although some studies found cubic or quadratic developmental trajectories of cortical thickness instead of a linear decline (e.g., Shaw et al., 2008; Raznahan et al., 2011), peaks always occurred before late childhood (8-10 years old) while after this period thickness only decreases, similar to our observation.

Secondly, cortical thickness of the LIFG predicted verbal memory scores in our sample. This is consistent with previous findings showing that there is a relationship between cortical thickness and verbal memory performance in adults (Walhovd et al., 2006; Dickerson et al., 2008), although the directionality of this relationship is different in our developmental study. In our sample, a general decrease in cortical thickness was associated with a general increase in verbal memory performance. In a sample with adults, individual differences in verbal memory abilities are positively correlated to measures of cortical thickness because such a sample is not characterised by maturational differences in cortical thinning. In another developmental study by Østby et al. (2011), focusing on visual memory instead of verbal memory (which has different developmental implications) found results similar to ours in which an increase in performance correlates with a decrease in cortical thickness. A developmental study most similar to our design (Sowell et al., 2001) already showed that an increase in verbal memory performance across development is associated with decreasing cortical thickness of the PFC. Since Sowell et al. (2001) used an intentional verbal learning task to test their hypotheses we could not dissociate whether PFC structural maturation truly affected mnemonic components of verbal memory development, or whether the effect was confounded by improved strategic abilities relying on the PFC as well. Based on our incidental, strategy-free encoding task, we can discard the latter possibility and conclude that PFC thinning most probably aids the mnemonic processes themselves.

More importantly, besides the replication of the effect of age on cortical thickness and the general effect of cortical thickness on verbal memory scores, we also confirmed our main hypothesis that cortical thickness mediates the developmental effect of age group on verbal memory scores. This mediation has to our knowledge not explicitly been tested in the literature before, and indicates that the development of verbal memory abilities with age is at least partially driven by maturational differences in cortical thickness of the LIFG. Our interpretation of this is that the cortical thinning, which we assume to indirectly reflect synaptic pruning (Petanjek et al., 2011), indeed allows for more efficient processing in the LIFG. In our sample, this efficiency then appears to be beneficial for a verbal memory task requiring semantic control processes that rely on the LIFG (Bokde et al., 2001; Badre & Wagner, 2007).

Functional specialisation

Another level of neural maturation that potentially underlies changes in mnemonic abilities over the course of development, is that of functional connectivity between relevant brain regions. We expected that due to protracted cortical thinning (i.e., assumed synaptic pruning) in the PFC, relevant changes in short- and long-range functional connectivity would occur within PFC regions during late childhood and adolescence (Jolles et al. 2011). This could for instance allow for specialisation of the IFG within larger functional networks. We hypothesised that in relation to a linguistic task as our verbal memory task, associated functional specialisation could be reflected in increased leftward lateralisation of relevant IFG – left MTL resting-state connectivity from childhood into early adulthood (Szaflarski et al., 2006 a; Holland et al., 2007). Therefore, we investigated whether increasingly left lateralized IFG – left MTL functional connectivity with increasing age group also mediates the effect of age group on verbal memory performance. This hypothesis could not be confirmed by our data. We did find an increase in leftward lateralisation of IFG – left MTL connectivity from childhood to adolescence, which is in line with previous studies, showing increasingly left lateralized task-related fMRI activity during development (reviewed in Holland et al., 2007). However, although mean laterality indices and mean memory scores develop along a similar trajectory, increased leftward lateralisation of IFG – left MTL connectivity with age did not significantly predict the increase in memory performance across development in our sample. This means that specialisation (i.e., lateralisation) of LIFG functional connectivity indeed increases with age group, but most likely does not mediate the increase in verbal memory performance with age.

Taken together, structural maturation of the LIFG appears to be important for verbal memory development, while there is no evidence that the development of left lateralized functional connectivity between the IFG and the left MTL influences the development of mnemonic abilities. This could mean that the maturation of the LIFG itself as a hub for semantic control specifically contributes to gains in performance with age, while the maturation of its interaction with the left MTL might not be of particular relevance. In accordance with the latter, we neither found an effect of mere LIFG – left MTL functional connectivity on verbal memory performance, although this maturation also showed a developmental trajectory similar to that of our memory scores. One argument to justify why only the specific maturation of the LIFG itself would matter for verbal memory development and not its interaction with the left MTL, would be that the MTL is assumed to play an important role in binding processes (Werkle-Bergner et al., 2006). Although most verbal memory studies report task-related MTL activity (see for a meta-analysis Kim, 2011), this region is more involved if the study material consists of verbal associations instead of single items. It might be the case that the relation between the design of our item-recognition task and the reliance on the left MTL is not strong enough to reach a significant association of LIFG – left MTL resting-state functional connectivity on task performance. Because this association is less direct than the task-related fMRI activity our hypotheses were based upon and would therefore profit from a more sensitive design (see 'Limitations'). This also means that specialisation or laterality of IFG – left MTL connectivity in particular might not have been the most optimal measure for our laterality index after all. Therefore, the current results on the laterality index do not rule out the possibility that specialisation of LIFG functional connectivity plays a role in increasing verbal memory performance with age group.

Another possibility to explain why we did find a mediation of LIFG cortical thickness maturation on the effect of age group on verbal memory scores, but not of lateralisation of IFG – left MTL functional connectivity, would be that the relationship between these two measures is not as strong as we assumed. Like previous studies (reviewed in Johnson, 2011), we assumed that cortical thinning and functional specialisation were driven by the same maturational mechanism, namely synaptic pruning leading to increased efficiency, and that both measures would therefore be highly negatively correlated. However, previous studies did not test this relationship directly. Interestingly, in our sample there was indeed a trend towards a negative correlation between cortical thickness of the LIFG and laterality indices of IFG – left MTL connectivity. Nonetheless, this relationship did not reach significance and thus does not provide us with conclusive evidence that it indeed exists. This means there is also a likelihood that cortical thinning and functional specialisation (i.e., lateralisation) during development might be driven by different underlying processes. The finding that cortical thinning mediates the effect of age group on memory score can also be due to other underlying processes than synaptic pruning, such as white matter myelination (Sowell et al., 2003). Intracortical myelination could impact measures of cortical thickness by changing the signal contrasts such that the boundary between white and gray matter, influencing the measure of cortical thickness, moves outward with increasing age (Mills & Tamnes, 2014). If myelination would indeed underlie the age and mediation effects we found for cortical thickness of the LIFG, this would also allow for more efficient processing in the LIFG, but does not necessarily affect its (lateralisation of) functional connectivity with the left MTL as was assumed for synaptic pruning (although it might).

Limitations

Laterality indices and memory scores seem to develop along a similar trajectory on a group level. However, we were unable to detect a significant effect of lateralisation of IFG – left MTL connectivity on memory performance across development in our sample. This could of course be due to a lack of such a relationship as we described above, but it is also possible that this relationship does exist. In the latter case, the design of our study was not sensitive enough to detect it. This could for instance be due to the distribution of our cross-sectional sample, which consists of three different age groups with large gaps in between. Hereby, information is lost because of the use of averages to create and compare group means, which is not optimal when investigating mediations. A continuous distribution however, in which all ages relevant for PFC maturation (i.e., between 10 and 32 years old) were represented, would have allowed us to truly model and compare neural and behavioural developmental trajectories. Another limitation of the use of a cross-sectional design is the amount of participants that is necessary to detect an effect of brain structure or function on behaviour (Steen, Hamer & Lieberman, 2007). In our cross-sectional study, inter-individual differences other than age contribute to the variation in our sample, which makes it more difficult to detect subtle maturational differences and effects than in longitudinal studies where there is no such variation. A future study would thus benefit from either a more continuous distribution of different age samples, or a longitudinal design. Thus, the most optimal design to investigate whether the maturation of neural structure and function mediates the development of verbal memory would be a longitudinal one, or a cross-sectional design with continuous sampling of different age groups and a substantial amount of subjects to represent each age bin. One valuable addition could also be a within-subjects contrast, for instance by examining both item recognition as well as an associative memory component. A primary reason for this would be that the assumed reliance on MTL structures is bigger for associative memory tasks, which would increase the chance of finding a relation between task performance and resting-state LIFG – left MTL connectivity. Besides this, such a within-subjects contrast would allow us to not only draw inferences on quantitative aspects of development, but also on a qualitative level: do different kinds of mnemonic representations develop differently from childhood into adulthood?

Conclusions

We aimed to investigate whether the maturation of neural structure and function is associated with behaviourally enhanced verbal memory performance from childhood into adulthood. In summary, we replicated previous studies showing an increase in verbal memory performance from childhood into adolescence and adulthood. Furthermore, we have shown that cortical thinning of the LIFG across the three age groups mediates the effect of age group on mnemonic abilities. This we interpreted as a reflection of the gains in semantic control efficiency of the LIFG after synaptic pruning has taken place during adolescence. There was no significant effect of lateralisation of IFG – left MTL functional connectivity on verbal memory performance. This indicates that structural LIFG maturation is indeed an important contributor to verbal memory development during adolescence, whereas functional LIFG specialisation at least seems to play a less significant role.

References

Alloway, T. P. (2007). *Automated Working: Memory Assessment: Manual.* Pearson.

Amlien, I.K., Fjell, A.M., Tamnes, C.K., Grydeland, H., Krogsrud, S.K., Chaplin, T.A., Rosa, M.G.P. & Walhovd, K.B. (2014). Organizing principles of human cortical development—thickness and area from 4 to 30 years: insights from comparative primate neuroanatomy. *Cerebral Cortex, 26,* 257-267.

Andersson, J.L.R., Jenkinson, M. & Smith, S. (2007a). Non-linear optimisation. *FMRIB technical report TR07JA1.*

Andersson, J.L.R., Jenkinson, M. & Smith, S. (2007b). Non-linear registration, aka Spatial normalisation. *FMRIB technical report TR07JA2*

Badre, D., & Wagner, A.D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45*(13), 2883-2901.

Baltes, P.B., Lindenberger, U. & Staudinger, U.M. (2006). Lifespan theory in developmental psychology. In: Damon, W. (Series Ed.), & Lerner, R.M. (Vol. Ed.), *Handbook of Child Psychology: vol. 1. Theoretical Models of Human Development,* 6. Wiley, New York.

Blachstein, H. & Vakil, E. (2016). Verbal learning across the lifespan: an analysis of the components of the learning curve. *Aging, Neuropsychology and Cognition, 23*(2), 133-153.

Brainerd, C.J. & Reyna, V.F. (1996). Mere memory testing creates false memories in children. *Developmental Psychology, 32,* 467–476.

Brainerd, C.J., Reyna, V.F., & Ceci, S.J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin, 134*(3), 343-382.

Brainerd, C.J., Reyna, V.F. & Kneer, R. (1995). False-recognition reversal: When similarity is distinctive. *Journal of Memory & Language, 34,* 157–185.

Bokde, A.L.W., Tagaments, M.A., Friedman, R.B. & Horwitz, B. (2001). Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. *Neuron, 30,* 609-617.

Clarkson, M.J., Cardoso, M.J., Ridgway, G.R., Modat, M., Leung, K.K., Rohrer, J.D., Fox, N.C. & Ourselin, S. (2011). A comparison of voxel and surface based cortical thickness estimation methods. *NeuroImage, 57*(3), 856-865.

Davis, H.P., Klebe, K.J., Guinther, P.M., Schroder, K.B., Cornwell, R.E. & James, L.E. (2013). Subjective Organization, Verbal Learning and Forgetting Across the Life Span: From 5 to 89. *Experimental Aging Research, 39,* 1-26.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, M.A., Maguire, R.P., Hyman, B.T., Albert, M.S. & Killiany, R.J. (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.

Dickerson, B.C., Fenstermacher, E., Salat, D.H., Wolk, D.A., Maguire, R.P., Desikan, R., Pacheco, J., Quinn, B.T., Van der Kouwe, A., Greve, D.N., Blacker, D., Albert, M.S., Killiany, R.J. & Fischl, B. (2008). Detection of cortical thickness correlates of cognitive performance: reliability across MRI scan sessions, scanners, and field strengths. *NeuroImage, 39*(1), 10-18.

Ducharme, S., Albaugh, M.D., Nguyen, T.V., Hudziak, J.J., Mateos-Pérez, J.M., Labbe, A., Evans, A.C., Karama, S. & Brain Development Cooperative Group. (2016). Trajectories of cortical thickness maturation in normal brain development—The importance of quality control procedures. *NeuroImage, 125,* 267-279.

Durston, S. & Casey, B.J. (2006). What have we learned about cognitive development from neuroimaging? *Neuropsychologia, 44*(11), 2149-2157.

Durston, S., Davidson, M.C., Tottenham, N., Galvan, A., Spicer, J., Fossella, J.A. & Casey, B.J. (2006). A shift from diffuse to focal cortical activity with development. *Developmental Science, 9*(1), 1-8.

Duverne, S., Motamedinia, S. & Rugg, M.D. (2009). The Relationship between Aging, Performance and the Neural Correlates of Successful Memory Encoding. *Cerebral Cortex, 19,* 733-744.

Everts, R., Lidzba, K., Wilke, M., Kiefer, C., Mordasini, M., Schroth, G., Perrig, W. & Steinlin, M. (2009). Strengthening of laterality of verbal and visuospatial functions during childhood and adolescence. *Human Brain Mapping, 30*(2), 473-483.

Fair, D.A., Dosenbach, N.U., Church, J.A., Cohen, A.L., Brahmbhatt, S., Miezin, F.M., Barch, D.M., Raichle, M.E., Petersenm S.E. & Schlaggar, B. L. (2007). Development of distinct control networks through segregation and integration. *Proceedings of the National Academy of Sciences, 104*(33), 13507-13512.

Fair, D.A., Cohen, A.L., Power, J.D., Dosenbach, N.U., Church, J.A., Miezin, F.M., Schlaggar, B.L. & Petersen, S.E. (2009). Functional brain networks develop from a "local to distributed" organization. *PLoS Computational Biology, 5*(5), e1000381.

Finn, A.S., Kalra, P.B., Goetz, C., Leonard, J.A., Sheridan, M.A. & Gabrieli, J.D.E. (2016). Developmental dissociation between the maturation of procedural memory and declarative memory. *Journal of Experimental Child Psychology, 142,* 212-220.

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of*

Sciences, 97(20), 11050-11055.

Fjell, A.M., Walhovd, K.B., Westlye, L.T., Østby, Y., Tamnes, C.K., Jernigan, T.L., Gamst, A. & Dale, A.M. (2010). When does brain aging accelerate? Dangers of quadratic fits in cross-sectional studies. *NeuroImage, 50*(4), 1376-1383.

Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences USA, 101*(21), 8174–8179.

Gogtay, N., Nugent, T.F., Herman, D.H., Ordonez, A., Greenstein, D., Hayashi, K.M. (2006). Dynamic mapping of normal human hippocampal development. *Hippocampus, 16*(8), 664–672.

Golby, A.J., Poldrack, R.A., Brewer, J.B., Spencer, D., Desmond, J.E., Aron, A.P. & Gabrieli, J.D.E. (2001). Material-specific lateralization in the medial temporal lobe and prefrontal cortex during memory encoding. *Brain, 124,* 1841-1854.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* Guilford Press.

Heaton, R.K., Chelune, G.J., Talley, J.L., Kay, G.G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test manual: Revised and expanded.* Odessa, FL: Psychological Assessment Resources.

Holland, S.K., Plante, E., Byars, A.W., Strawsburg, R.H., Schmithorst, V.J. & Ball, W.S. (2001). Normal fMRI brain activation patterns in children performing a verb generation task. *NeuroImage, 14*(4), 837-843.

Holland, S.K., Vannest, J., Mecoli, M., Jacola, L.M., Tillema, J.M., Karunanayaka, P.R., Schmithorst, V.J., Yuan, W., Plante, E. & Byars, A. W. (2007). Functional MRI of language lateralization during development in children. *International Journal of Audiology, 46*(9), 533-551.

Hutton, C., Draganski, B., Ashburner, J. & Weiskopf, N. (2009). A comparison between voxel-based cortical thickness and voxel-based morphometry in normal aging. *NeuroImage, 48*(2), 371-380.

Jenkinson, M., Bannister, P., Brady, M. & Smith, S. (2002). Improved Optimisation for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage, 17*(2), 825-841.

Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W. & Smith, S.M. (2012). FSL. *NeuroImage, 62,* 782-790.

Johnson, M.H. (2001). Functional brain development in humans. *Nature Reviews Neuroscience, 2,* 475-483.

Johnson, M.H. (2011). Interactive specialization: a domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience, 1*(1), 7-21.

Jolles, D.D., Van Buchem, M.A., Crone, E.A. & Rombouts, S.A. (2011). A comprehensive study of whole-brain functional connectivity in children and young adults. *Cerebral Cortex, 21*(2), 385-391.

Kelly, A.C., Di Martino, A., Uddin, L.Q., Shehzad, Z., Gee, D.G., Reiss, P.T., Margulies, D.F., Castellanos, F.X. & Milham, M.P. (2009). Development of anterior cingulate functional connectivity from late childhood to early adulthood. *Cerebral Cortex, 19*(3), 640-657.

Kim, H. (2011). Neural activity that predicts subsequent memory and forgetting: A meta-analysis of 74 fMRI studies. *NeuroImage, 54,* 2446-2461.

Koolschijn, P.C.M. & Crone, E.A. (2013). Sex differences and structural brain maturation from childhood to early adulthood. *Developmental Cognitive Neuroscience, 5,* 106-118.

Lenroot, R.K. & Giedd, J.N. (2006). Brain development and adolescents: insights from anatomical magnetic resonance imaging. *Neuroscience and Biobehavioral Reviews, 30,* 718-729.

Liu, X., Qin, S., Rijpkema, M., Luo, J. & Fernández, G. (2010). Intermediate Levels of Hippocampal Activity Appear Optimal for Associative Memory Formation. *PLoS ONE, 5*(10), e13147.

MacMillan, N.A. & Creelman, C.D. (2005). *Detection Theory: A User's Guide.* Psychology Press, 7.

Maril, A., Avital, R., Reggev, N., Zuckerman, M., Sadeh, T., Ben Sira, L., et al. (2011). Event congruency and episodic encoding: a developmental fMRI study. *Neuropsychologia, 49*(11), 3036–3045.

Mills, K.L., & Tamnes, C.K. (2014). Methods and considerations for longitudinal structural brain imaging analysis across development. *Developmental Cognitive Neuroscience, 9,* 172-190.

Müller et al. (in preparation).

Nielsen, J.A., Zielinski, B.A., Ferguson, M.A., Lainhart, J.E. & Anderson, J.S. (2013). An Evaluation of the Left-Brain vs. Right-Brain Hypothesis with Resting State Functional Connectivity Magnetic Resonance Imaging. *PLoS ONE, 8*(8), e71275.

Ofen, N. & Shing, Y.L. (2013). From perception to memory: Changes in memory systems across the lifespan. *Neuroscience and Biobehavioral Reviews, 37*(9B), 2258–2267.

Opitz, B., Mecklinger, A. & Friederici, A.D. (2000). Functional asymmetry of human prefrontal cortex: Encoding and retrieval of verbally and nonverbally coded information. *Learning and Memory, 7,* 85-96.

Østby, Y., Tamnes, C.K., Fjell, A.M., & Walhovd, K.B. (2011). Dissociating memory processes in the developing brain: the role of hippocampal volume and cortical thickness in recall after minutes versus days. *Cerebral Cortex, 22,* 381-390.

Petanjek, Z., Judaš, M., Šimić, G., Rašin, M.R., Uylings, H.B., Rakic, P. & Kostović, I. (2011). Extraordinary neoteny of synaptic spines in the human prefrontal cortex. *Proceedings of the National Academy of Sciences, 108*(32), 13281-13286.

Pruim, R.H.R., Mennes, M., Van Rooij, D., Llera, A., Buitelaar, J.K. & Beckmann, C.F. (2015). ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *NeuroImage, 112,* 267-277.

Qin, S., Rijpkema, M., Tendolkar, I., Piekema, C., Hermans, E. J., Binder, M., Petersson, K.M., Luo, J. & Fernández, G. (2009). Dissecting medial temporal lobe contributions to item and associative memory formation. *NeuroImage, 46*(3), 874-881.

Ragland, J.D., Turetsky, B.I., Gur, R.C., Gunning-Dixon, F., Turner, T., Schroeder, L., Chan, R. & Gur, R. E. (2002). Working memory for complex figures: an fMRI comparison of letter and fractal n-back tasks. *Neuropsychology, 16*(3), 370-379.

Raz, N. & Rodrigue, K.M. (2006). Differential aging of the brain: patterns, cognitive correlates and modifiers. *Neuroscience and Biobehavioral Reviews, 30,* 730-748.

Raznahan, A., Shaw, P., Lalonde, F., Stockman, M., Wallace, G.L., Greenstein, D., Clasen, L., Gogtay, N. & Giedd, J.N. (2011). How does your cortex grow? *The Journal of Neuroscience, 31*(19), 7174-7177.

Ressel, V., Wilke, M., Lidzba, K., Lutzenberger, W. & Krägeloh-Mann, I. (2008). Increases in language lateralization in normal children as observed using magnetoencephalography. *Brain and Language, 106*(3), 167-176.

Ségonne, F., Pacheco, J. & Fischl, B. (2007). Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *Medical Imaging, IEEE Transactions on, 26*(4), 518-529.

Shaw, P., Kabani, N.J., Lerch, J.P., Eckstrand, K., Lenroot, R., Gogtay, N., Greenstein, D., Clasen, L., Evans, A., Rapoport, J.L., Giedd, J.N. & Wise, S.P. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *The Journal of Neuroscience, 28*(14), 3586-3594.

Shing, Y.L., Werkle-Bergner, M., Brehmer, Y., Müller, V., Li, S.C. & Lindenberger, U. (2010). Episodic memory across the lifespan: The contributions of associative and strategic components. *Neuroscience & Biobehavioral Reviews, 34*(7), 1080-1091.

Shing, Y.L., Werkle-Bergner, M., Li, S.C. & Lindenberger, U. (2008). Associative and strategic components of episodic memory: A life-span dissociation. *Journal of Experimental Psychology General, 137,* 495-513.

Shrout, P.E. & Bolger, N. (2002). Mediation in Experimental and Nonexperimental Studies: New Procedures and Recommendations. *Psychological Methods, 7*(4), 422-445.

Simons, J.S. & Spiers, H.J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience, 4,* 637-648.

Sled, J.G., Zijdenbos, A.P. & Evans, A.C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *Medical Imaging, IEEE Transactions on, 17*(1), 87-97.

Smith, S. (2002). Fast Robust Automated Brain Extraction. *Human Brain Mapping, 17*(3), 143-155.

Sowell, E.R., Delis, D., Stiles, J., & Jernigan, T.L. (2001). Improved memory functioning and frontal lobe maturation between childhood and adolescence: a structural MRI study. *Journal of the International Neuropsychological Society, 7*(3), 312-322.

Sowell, E.R., Peterson, B.S., Thompson, P.M., Welcome, S.E., Henkenius, A.L. & Toga, A.W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience, 6,* 309-315.

Sowell, E.R., Thompson, P.M., Leonard, C.M., Welcome, S.E., Kan, E., Toga, A.W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *Journal for Neuroscience, 24*(38), 8223–8231.

Staresina, B.P., Duncan, K.D., Davachi, L. (2011). Perirhinal and parahippocampal cortices differentially contribute to later recollection of object- and scene-relatedevent details. *Journal for Neuroscience, 31*(24), 8739–8747.

Steen, R.G., Hamer, R.M., & Lieberman, J.A. (2007). Measuring brain volume by MR imaging: impact of measurement precision and natural variation on sample size requirements. *American Journal of Neuroradiology, 28*(6), 1119-1125.

Supekar, K., Musen, M. & Menon, V. (2009). Development of large-scale functional brain networks in children. *PLoS Biology, 7*(7), e1000157.

Szaflarski, J.P., Holland, S.K., Schmithorst, V.J., & Byars, A.W. (2006a). fMRI study of language lateralization in children and adults. *Human Brain Mapping, 27*(3), 202-212.

Szaflarski, J.P., Schmithorst, V.J., Altaye, M., Byars, A.W., Ret, J., Plante, E. & Holland, S. K. (2006b). A longitudinal functional magnetic resonance imaging study of language development in children 5 to 11 years old. *Annals of Neurology, 59*(5), 796-807.

Tavor, I., Jones, O.P., Mars, R.B., Smith, S.M., Behrens, T.E. & Jbabdi, S. (2016). Task-free MRI predicts individual differences in brain activity during task performance. *Science, 352*(6282), 216-220.

Thompson-Schill, S.L., D'Esposito, M., Aguirre, G.K. & Farah, M.J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proceedings of the National Academy of Sciences USA, 94,* 14792-14797.

Uddin, L.Q., Supekar, K. & Menon, V. (2010). Typical and atypical development of functional human brain networks: insights from resting-state FMRI. *Frontiers in Systems Neuroscience, 4,* 21.

Vakil, E. & Blachstein, H. (1997). Rey ALVT: Developmental norms for adults and the sensitivity of different memory measures to age. *The Clinical Neuropsychologist, 11,* 356-369.

Vakil, E., Blachstein, H. & Sheinman, M. (1998). Rey ALVT: Developmental norms for children and the sensitivity of different memory measures to age. *Child Neuropsychology, 15,* 21-39.

Wagner, A.D., Schacter, D.L., Rotte, M., Koutstaal, W., Maril, A., Dale, A.M., Rosen, B.R. & Buckner, R.L. (1998). Building Memories: Remembering and Forgetting of Verbal Experiences as Predicted by Brain Activity. *Science, 281,* 1188-1191.

Walhovd, K.B., Fjell, A.M., Dale, A.M., Fischl, B., Quinn, B.T., Makris, N., Salat, D. & Reinvang, I. (2006). Regional cortical thickness matters in recall after months more than minutes. *NeuroImage, 31*(3), 1343-1351.

Werkle-Bergner, M., Müller, V., Li, S.C. & Lindenberger, U. (2006). Cortical EEG correlates of succesful memory encoding: Implications for lifespan comparisons. *Neuroscience and Biobehavioral Reviews, 30,* 839-854.

Wierenga, L.M., Langen, M., Oranje, B. & Durston, S. (2014). Unique developmental trajectories of cortical thickness and surface area. *NeuroImage, 87,* 120-126.

Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., Fischl, B., Liu, H. & Buckner, R.L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology, 106*(3), 1125-1165.