



Universiteit
Leiden
The Netherlands

The adolescent brain : unraveling the neural mechanisms of cognitive and affective development

Peters, S.

Citation

Peters, S. (2016, January 27). *The adolescent brain : unraveling the neural mechanisms of cognitive and affective development*. Retrieved from <https://hdl.handle.net/1887/37391>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/37391>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/37391> holds various files of this Leiden University dissertation

Author: Peters, Sabine

Title: The adolescent brain : unraveling the neural mechanisms of cognitive and affective development

Issue Date: 2016-01-27

Chapter 6

Longitudinal development of the frontoparietal network: Contributions of age, performance, working memory and brain structure



This chapter is based on:

Peters, S., Van Duijvenvoorde, A.C.K., P.C.M.P. Koolschijn & Crone, E.A. Longitudinal development of neural activity in the frontoparietal network: Contributions of age, performance, working memory and brain structure (in revision, 2015).

Abstract

Even though it is well conceptualized that neural activity in the frontoparietal network changes during childhood and adolescent development, there is surprisingly little consensus about the direction of change, and a comprehensive study is lacking. Using a large-scale longitudinal fMRI study, we aimed to test growth patterns across childhood and adolescence in frontoparietal activity during a feedback learning task capturing multiple aspects of cognitive control. Our first aim was to test for linear and non-linear developmental trajectories of activity in the dorsolateral prefrontal cortex (DLPFC), superior parietal cortex (SPC) and the pre-supplementary motor area/anterior cingulate cortex (pre-SMA/ACC). Our second aim was to test which factors drive developmental change in the frontoparietal network besides age. Contributions of task performance, working memory and cortical thickness were investigated. To these ends, a developmental sample ($N = 208$, 8-27 years old) was tested twice across a period of two years. The results showed that developmental patterns for neural activity in DLPFC and SPC were best characterized by a quadratic age function leveling off/peaking in late adolescence, and by a linear increase across age in pre-SMA/ACC. In addition to age, task performance explained variance in DLPFC and SPC activity, but not in pre-SMA/ACC activity. In contrast, cortical thickness explained additional variance in pre-SMA/ACC activity, but not in DLPFC and SPC. Together, these findings provide a novel perspective of developmental changes in the frontoparietal network, arguing against a simple imbalance model with linear development of cognitive control regions.

Introduction

Although adolescent development of neural activity during cognitive control is studied extensively, results from cross-sectional studies remain inconclusive. A meta-analysis (Crone & Dahl, 2012) showed that, although many fMRI studies reported age-related increases in prefrontal and parietal recruitment during cognitive tasks, a substantial number of studies reported age-related decreases in these brain regions. Until recently most studies only used cross-sectional comparisons to characterize development, which has several disadvantages compared to longitudinal designs. Here, we used a large longitudinal dataset to test growth patterns across adolescent development. Specifically, we tested whether frontoparietal activity follows a linear pattern (i.e. monotonic development over time, no adolescent-specific changes), a quadratic pattern (i.e., adolescent-specific effects) or a cubic pattern (adolescent-emergent; e.g. stable levels during childhood, steep changes in adolescence and stabilization in adulthood) (Braams et al., 2015; Somerville et al., 2013). A second goal was to test which other factors influence changes over time in frontoparietal activity in addition to age. Here we tested the contributions of increased task performance, working memory capacity and cortical thickness.

Cognitive development has been hypothesized to follow a linear developmental trajectory according to ‘dual-systems models’ of adolescent brain development, with steadily increasing frontoparietal recruitment from childhood to adulthood (Ernst et al., 2006; Somerville & Casey, 2010; Steinberg, 2008). Several authors have argued for adaptations to dual-systems models, by suggesting that the frontoparietal network is not necessarily immature or inaccessible in adolescents, but sensitive to different situations than adults depending on e.g. motivational salience, or increased specialization of brain areas for more specific tasks (Casey, 2015; Crone & Dahl, 2012; Johnson, 2011; Pfeifer & Allen, 2012). This is for instance supported by results from feedback learning paradigms which demonstrated that adults showed increased recruitment of prefrontal and parietal cortex following negative compared to positive feedback, whereas children showed more activity in these same regions for positive compared to negative feedback (Peters, Braams, et al., 2014; van den Bos et al., 2009; van Duijvenvoorde et al., 2008). These findings indicate that young adolescents are capable of recruiting frontoparietal regions but in different situations than adults, arguing against a simple frontoparietal immaturity model with linear development in cognitive control regions.

The current study builds on prior cross-sectional findings by testing for linear and non-linear development in neural activity during feedback learning using an experimental paradigm inspired by the Wisconsin Card Sorting Test (Milner, 1963). We designed this task to capture multiple aspects of cognitive control functioning and we focused on three regions within the frontoparietal network, dorsolateral prefrontal cortex (DLPFC), pre-supplementary motor area/anterior cingulate cortex (pre-SMA/ACC) and superior parietal cortex (SPC). Prior studies revealed that these are key regions for cognitive control, as demonstrated by meta-analyses across

multiple executive functions (Kim et al., 2012; Niendam et al., 2012). In addition, our rationale for focusing on these three regions is based on findings showing that DLPFC, pre-SMA/ACC and SPC show pronounced age-related changes in activity during cognitive control tasks such as working memory (Klingberg et al., 2002; Thomason et al., 2009) and feedback learning paradigms (Crone et al., 2008; Peters, Braams et al., 2014; van Duijvenvoorde et al., 2008). However, to date little is known about how activity in the frontoparietal network changes longitudinally.

Longitudinal designs have critical advantages over cross-sectional designs. For instance, previous studies demonstrated important individual differences in developmental trajectories that can be overlooked in cross-sectional designs (Koolschijn et al., 2011; Ordaz, Foran, Velanova, & Luna, 2013; Shaw et al., 2013). Furthermore, longitudinal designs have increased power to detect developmental change, because testing within-individual changes reduces error related to cohort differences (Fjell et al., 2010; Koolschijn et al., 2011).

Besides investigating age-related patterns of neural activity, a second goal of this study was to investigate other factors influencing time-related changes in frontoparietal activity in addition to age. There are multiple processes closely related to advancing age that may drive changes in neural activity. That is, an increase in age could be the only factor explaining time-related increases or decreases in activity, but other factors might also play a role. The factors investigated in this study were task performance, working memory capacity and structural brain development. Task performance has been shown to influence neural activity, and there is evidence that a portion of developmental changes attributed to advancing age are related more to increases in performance (Church et al., 2010; Koolschijn et al., 2011). Here we tested whether performance on the feedback-learning task partly explained changes in neural activation over time. Working memory has previously been argued to be a core prerequisite for cognitive development (Case, 1992) and cognitive control functions (Huizinga et al., 2006), and as such was investigated as an important contributor to changes over time in neural activity during feedback-learning. That is, we aimed to study whether a portion of changes in neural activity during feedback learning was explained by individual differences in working memory capacity. A final factor that was investigated is cortical thickness. Several cross-sectional studies have suggested a link between functional activity and structural gray matter in adults (Harms, Wang, Csernansky, & Barch, 2013; Hegarty et al., 2012) and children (Lu et al., 2009; Wendelken, O'Hare, Whitaker, Ferrer, & Bunge, 2011). It is likely that developmental changes in neural activity are at least partly influenced by structural development of these brain regions, although the longitudinal relation between structural maturation and development of brain function is not well understood.

Taken together, in this study, we tested developmental trajectories of activation in the frontoparietal network in a large longitudinal fMRI sample across a wide age range ($N = 208$, 8-27 years) with a two year interval between the first and second time point. Our aims were 1) to examine growth trajectories of core areas in the frontoparietal network (DLPFC, pre-SMA/ACC and SPC) and to define the shape of age-related changes, 2) to test the additional contributions of task

performance, working memory capacity and structural development to changes over time in neural activity.

Methods

Participants

At time point 1 (TP1), a total of 299 participants between ages 8-27 years underwent an MRI scan, of which 293 participants completed the feedback learning task in the MRI scanner. Of these, 25 participants were excluded from further analyses because of excessive movement (movement > 3.0 mm: $n = 19$), artifacts ($n = 3$) or because they were extreme outliers in task performance (> 3x the interquartile range: $n = 3$). In total, 268 participants were included at TP1 (*Mean Age* = 14.52 years, *SD* = 3.55; published in Peters, Braams et al., 2014). At time point 2 (TP2), a total of 254 of the initial 299 participants were scanned again approximately two years later (mean time = 1.99 years, *SD* = 0.10 years, range = 1.66-2.47 years). Reasons for not collecting a scan at TP2 ($n = 45$) were braces ($n = 32$) or no interest in participating again ($n = 13$). Further exclusions at TP2 were because of excessive movement at TP2 ($n = 9$), scanner artifacts ($n = 5$), loss of signal ($n = 3$) or extreme outliers (> 3x the interquartile range) on task performance ($n = 2$).

Only those participants who were included at both TP1 and at TP2 were included in the analyses ($N = 208$). All analyses were performed on these 208 participants, except for the analyses including working memory capacity and cortical thickness. For working memory, data were incomplete for five participants at TP1 and for two participants at TP2. For the analyses involving structural MRI data, visual quality control led to exclusion of 28 out of 208 participants: Three exclusions for insufficient quality data at both TP1 and TP2, 16 for TP1 and nine for TP2. These participants were only excluded from the analyses where cortical thickness was a factor. Taken together, the analyses with fMRI in the model contained a total of 208 participants (105 females), the analyses with working memory a total of 201 participants and the analyses with structural MRI in the model contained a total of 177 participants.

IQ was estimated with two subtests of the WAIS-III or WISC-III (Similarities and Block Design at TP1, Vocabulary and Picture Completion at TP2). The estimated IQ-scores of the 208 included participants were within the normal range at TP1 (85-143, *Mean* = 110.91, *SD* = 9.74) and TP2 (80-147, *Mean* = 108.92, *SD* = 10.18). The study was approved by the Institutional Review Board at the Leiden University Medical Center and all participants (or participant's parents in case of minors) provided written informed consent. Adults received payment for participation and children and their parents received small presents and payment for participation. Participants did not report psychiatric or neurological diagnoses, and no current use of psychotropic medication. All anatomical MRI scans were reviewed and cleared by a radiologist.

Feedback Learning Task

Participants performed a child-friendly feedback learning task in the MRI scanner described in detail earlier (Peters, Braams, et al., 2014; Peters, Koolschijn, Crone, Van Duijvenvoorde, & Raijmakers, 2014). In short, on each trial, participants viewed a screen with three boxes at the top part of the screen (Figure 1a). At the bottom part of the screen, a stimulus picture was presented, which was one of three possible stimuli. Participants were informed that all pictures belonged in one of the three boxes and that they had to find the correct box for each picture. Performance feedback was provided in the form of a plus-sign ('+') for correct choices (positive feedback) and a minus-sign ('-') for incorrect choices (negative feedback). Stimuli were presented in a pseudorandom order (maximum two identical pictures in a row). The sequence ended after 12 trials, or when participants chose the correct location twice for all three stimuli. Subsequently, a new sequence with three new pictures was presented. In total, participants completed 15 sequences, which resulted in a maximum of 180 trials. The task was divided into two blocks of eight and seven sequences, respectively. Before performing the task in the MRI scanner, participants practiced three sequences in a separate practice session. All trials started with a 500 ms fixation cross, followed by a 2500 ms time window during which the stimulus was presented and a response needed to be given. Feedback was presented for 1000ms. Inter-trial intervals were jittered with OptSeq (Dale, 1999), with intervals between 0-6 seconds in addition to the 500 ms fixation cross.

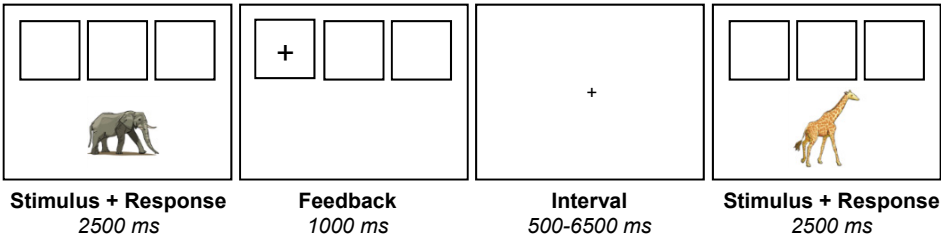


Figure 1a: Display of task sequence for the feedback learning task. During the last 500 ms of the Interval screen, a fixation cross was presented to prepare the participant for the next upcoming stimulus.

Feedback types

We distinguished between a learning phase and an application phase for all stimuli. The learning phase was defined as the trials where participants had not yet found the correct location for each stimulus and were still using feedback to find the correct locations. The application phase was defined as the trials where each stimulus was sorted correctly previously and was continued to be sorted correctly on subsequent trials. We excluded trials in the learning phase that did not result in learning, i.e., the trials where the feedback was not successfully used on the subsequent trial

(5.71 % of the trials). Based on the learning-application distinction, we defined the following three feedback types: Learning phase. a) Positive Learning: A first correct feedback for a stimulus followed by a correct sort on the next trial for this stimulus. b) Negative Learning: A first encountered incorrect feedback for a stimulus followed by a choice for another location on the next trial of this stimulus. Application phase. c) Application: Correct (i.e., positive) feedback for a stimulus that was sorted correctly before.

As a task performance measure, we calculated the ‘learning rate’ for each participant. This was defined as the percentage of trials in the learning phase for which feedback was successfully used on the next trial, compared to the total number of trials during the learning phase (including trials which did not result in learning according to the participants’ behavior on the next trial).

Working Memory task

Working memory was assessed with the Mental Counters task (Larson, Merritt, & Williams, 1988), which has been shown in a prior developmental study to be a well-suited task to measure the latent factor of working memory in children (Huizinga et al., 2006). Similar to the feedback learning paradigm, the mental counters task has both a spatial aspect (because two counters at two different locations have to be remembered) and a verbal rehearsal component, allowing us to assess the working memory aspects of the feedback learning task contributing to neural activity. Participants were presented with a screen with two horizontal lines (the ‘counters’) placed next to each other (Figure 1b).

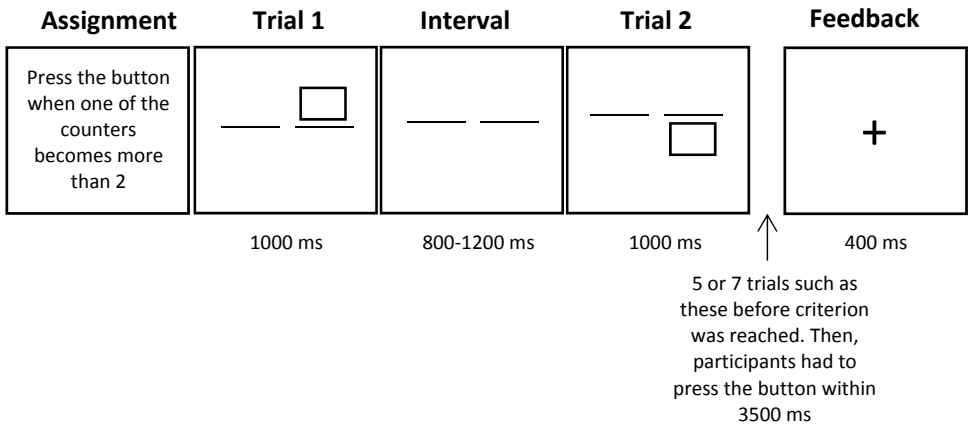


Figure 1b: Display of task sequence for the Mental Counters task as a measure of working memory capacity.

At each trial, a square randomly appeared on top of or below one of the two horizontal lines. The participant was instructed to keep track of the ‘score’ of the two counters. The value of the counters changed on each trial: e.g., a square appearing above the left counter changed the score to 1-0. If on a next trial a square appeared above the right counter, the score changed to 1-1. A square appearing below one of the counters meant a point had to be subtracted for that counter. Participants were instructed to press a button when one of the counters reached a certain criterion value, e.g. ‘press when the score for one of the counters reaches more than two points’, and the criterion changed for each series. The amount of trials before criterion was reached was set to either five or seven trials. In total, 16 series were presented. Trials were separated by 800-1200 ms intervals and participants had a time window of 3500 ms to respond once criterion was reached. Feedback was provided as a ‘+’ for a correct button press, a ‘-’ for an incorrect button press and an ‘x’ for an omission. Performance on the working memory task was defined as the proportion of correct responses (at TP1: $M = 0.83$, $SD = 0.15$; at TP2: $M = 0.87$, $SD = .12$).

FMRI Data Acquisition

We used the same Philips 3.0 Tesla MRI scanner and settings for both time points (Peters, Braams, et al., 2014; Peters, Koolschijn, et al., 2014). Functional scans were acquired with T2*-weighted echo-planar imaging, for which the first two volumes were discarded to allow for equilibration of T1 saturation effects. The following settings were used: TR = 2.2 s, TE = 30 ms, sequential acquisition, 38 slices, slice thickness = 2.75 mm, Field of View (FOV) = 220 × 220 × 114.68 mm. We acquired a high-resolution 3D T1-FFE anatomical scan after the experimental task (TR = 9.76 ms, TE = 4.59 ms, 140 slices, voxel size = 0.875 mm, FOV = 224 × 177 × 168 mm). The experimental task was projected on a screen that was viewed through a mirror attached to the head coil. Participants were accustomed to the MRI environment and sounds with a mock scanner prior to the actual MRI scan.

FMRI Data Analysis

We performed two types of analyses: a whole-brain analysis for an illustrative overview of brain activity at TP1 and TP2, and regions-of interest (ROI) analyses for growth curve modeling. For all analyses we used SPM8 (Wellcome Department of Cognitive Neurology, London) to analyze fMRI data. All scans were corrected for slice timing acquisition and rigid body motion. All volumes were spatially normalized to T1 templates, using a 12-parameter affine transform with a nonlinear transformation involving cosine basis functions with resampling of the volumes to 3 mm voxels. T1 templates were based on the MNI305 stereotaxic space (Cocosco et al., 1997), an approximation of Talairach space (Talairach & Tournoux, 1988). Functional volumes were spatially smoothed with an 8 mm FWHM isotropic Gaussian kernel. The fMRI time series data were modeled by convolving a series of events with a hemodynamic response function. The modeled feedback events were categorized as: “Positive Learning”, “Negative Learning”, and “Applica-

tion”, which were time-locked with 0-duration to the moment of feedback presentation. All other trials (e.g., trials that did not result in learning or too-late trials) were modeled as events of no interest. These events were used as covariates in a general linear model together with a set of cosine functions that high-pass filtered the data. The least-squares parameter estimates of height of the best-fitting canonical HRF for each condition were used in pair-wise contrasts.

The main fMRI contrast was the learning contrast: Learning (Positive and Negative combined) > Application. With this contrast, we investigated neural activity for feedback that is informative for learning, averaged across negative and positive valence, relative to application. The contrast images were submitted to higher-level group analyses. Whole-brain fMRI analyses were performed with an FWE-corrected threshold at $p < .05$.

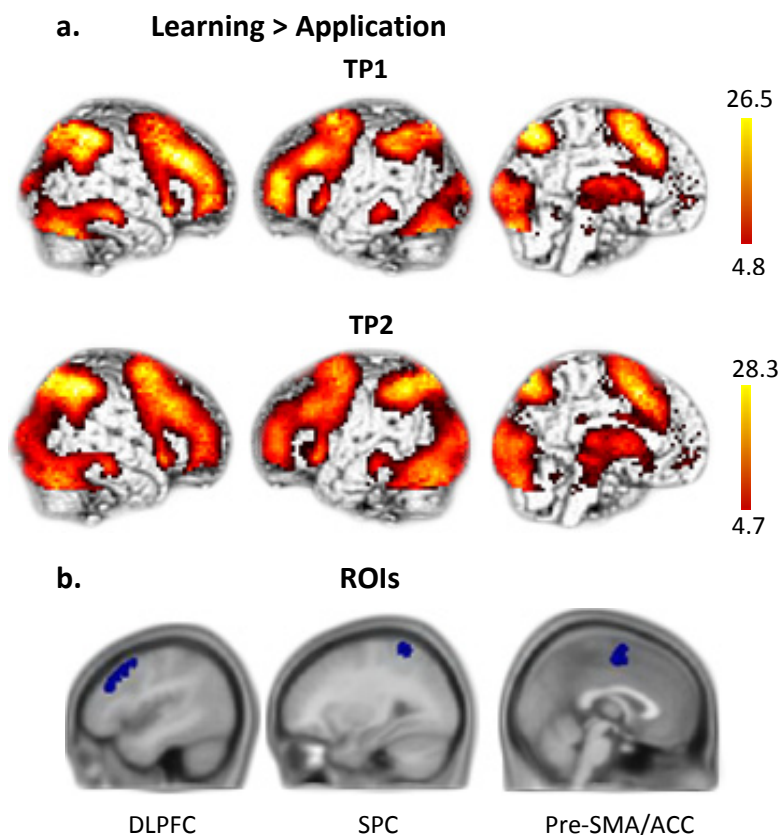


Figure 2a: Whole-brain analyses showing comparable neural activation patterns at TP1 and TP2 (FWE-corrected at $p < .05$). 2b: Bilateral Regions of Interest in the DLPFC, SPC and pre-SMA/ACC, extracted from the Harvard-Oxford Cortical Atlas. The anatomical ROIs were situated within the activation maps of the Learning > Application contrast.

Region-of-Interest analyses

Region-of-interest (ROI) analyses were performed with the MarsBaR toolbox (v. 0.42) in SPM8 (Brett et al., 2002). We used bilateral anatomical ROIs which were obtained from the probabilistic Harvard-Oxford Cortical Structural atlas. The ROIs were Middle Frontal Gyrus for DLPFC, Superior Parietal Lobule for SPC, and Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex) for pre-SMA/ACC. Because the anatomical ROIs based on probability maps were large and to ensure the selected voxels had a high probability of belonging to the targeted ROI, we created more focal ROIs by thresholding at 50 %, indicating that for each voxel, the probability that the voxel was actually part of the ROI (e.g. DLPFC) was > 50 % (see Figure 2b). These anatomical ROIs fell within activity clusters for the contrast Learning > Application (FWE-corrected at $p < .05$) (see Figure 2a). Beta values reflecting activity for all voxels within each ROI were averaged to produce a mean signal for each ROI per time point.

Structural brain analysis

Cortical reconstruction was measured automatically using FreeSurfer 5.3 (<http://surfer.nmr.mgh.harvard.edu>). Structural brain maturation of TP1 was reported earlier using a prior version of FreeSurfer (Koolschijn, Peper, & Crone, 2014), therefore we reconstructed and reanalyzed all anatomical scans from the first wave using version 5.3. Details of the surface-based cortical reconstruction and subcortical volumetric segmentation procedures have been extensively documented previously (Dale et al., 1999; Fischl & Dale, 2000; Ségonne et al., 2004).

To extract reliable volume and thickness estimates, images were automatically processed with the longitudinal stream in FreeSurfer (Reuter, Schmansky, Rosas, & Fischl, 2012). Specifically, an unbiased within-subject template space and image (Reuter & Fischl, 2011) is created using robust, inverse consistent registration (Reuter, Rosas, & Fischl, 2010). Several processing steps, such as skull stripping, Talairach transforms, atlas registration as well as spherical surface maps and parcellations were then initialized with common information from the within-subject template, significantly increasing reliability and statistical power (Reuter et al., 2012).

To extract average cortical thickness from the FSL anatomical ROIs, we performed the following steps: 1) Each anatomical ROI (DLPFC, pre-SMA/ACC and SPC) was registered automatically to the FreeSurfer “fsaverage” template with normalized mutual information and inspected for accuracy of registration. Of note, as FreeSurfer calculates cortical thickness per hemisphere, the ROIs were split into a left and right structural ROI. 2) Individual cortical thickness data was mapped to the “fsaverage” template. 3) Average cortical thickness in mm was extracted for each ROI and individual separately. 4) For subsequent analyses, all bilateral ROIs, including pre-SMA/ACC, were averaged to a single ROI as results were comparable between hemispheres. We used an average weighted procedure by taking into account hemispherical differences in surface size maps.

Statistical analyses

To model the shape of individual growth curves, we used mixed model analyses (also termed “random effects”, “multilevel modeling”, or hierarchical linear model-analyses) on the ROI values for contrasts of interest (Ordaz et al., 2013). This method expands on multiple regression analyses and is suited for longitudinal data because it takes into account the repeated-nature of the data, and controls for the dependency in measures within individuals (i.e., nested data). It was not necessary to calculate change scores, because mixed models take into account all data including individual differences in intercepts. We performed mixed analyses with the NLME package in R (Pinheiro, Bates, DebRoy, & Sarkar, 2007) version 3.1.0. With this package it is possible to test for fixed effects (effects that are similar for all participants) and random effects (effects that vary across participants) of age on brain activity. Models were compared using the Akaike Information Criterion (AIC), a standard measure for model comparison which indicates how well the model describes the data. Lower AIC values indicate a better fit of the model to the data. For nested models (i.e., comparing models with only 1 different term), we additionally tested with log-likelihood tests (χ^2) whether changes in model fit were significant. Our goals were twofold: 1) to test which shape of age (linear, quadratic cubic) best described the developmental pattern for the variables neural activity, task performance, working memory capacity and cortical thickness, and 2) to investigate which factors explain variance in brain activity above age; task performance, working memory capacity and/or cortical thickness. The model-building steps are described in the next paragraphs.

Developmental patterns: linear, quadratic and cubic trajectories

We first tested for each dependent variable (task performance, working memory capacity and activity and cortical thickness for each ROI) which age shape best described the developmental pattern. The base-model included a fixed intercept and a random intercept, with the latter capturing the variation in the intercept to account for the repeated nature of the data. Next, the base-model was tested against three models that tested the shape of the grand mean trajectory for age. We tested for a linear effect of age (i.e., monotonic development), a quadratic effect of age (i.e., an adolescent-specific effect) and a cubic effect (i.e., adolescent-emergent pattern) by adding three polynomial functions for age to the base-model (Braams et al., 2015; Somerville et al., 2013). We only selected a linear, quadratic or cubic model if the age term resulted in a better fit compared to the base-model without age as indicated by the AIC and a log-likelihood test. For the best model, we tested whether specifying age as an effect with a random slope resulted in a better fit (judged by the AIC and a log-likelihood test) compared to age as a fixed effect. A significant random slope would indicate that the effect of age on the dependent measure differs for each individual. An example of a formal notation of such a mixed model to predict, for instance, task performance, would be as follows:

$$\begin{aligned} \text{Performance} &= \pi_{0i} + \pi_{1i} (\text{Age})_i + e_{ii} \\ \text{With} \quad \pi_{0i} &= \beta_{00} + r_{0i} \\ \pi_{1i} &= \beta_{10} + r_{1i} \end{aligned}$$

In this example model, substitution of the second level model into the first level model gives the integrated model that was fitted to the data. β_{00} reflects the grand mean intercept of neural activity at the average age of the sample. β_{10} reflects the grand mean slope of age effects on neural activity, and e_{ii} represents the residual error term. This example model displays a random intercept (r_{0i}) indicating different starting points of development for each participant, and a random slope of age (r_{1i}), indicating individual variability in the change over time. All models were fit with full information maximum likelihood estimates. A random slope did not improve model fit in any of the analyses we performed, except for the quadratic age effect on working memory capacity ($p < .001$). Therefore, we did not describe this further in the results section.

Explaining development of neural activity with different predictors

The second aim was to investigate which factors explain development of neural activity in the frontoparietal network in addition to age. The predictors we tested for significance above age were task performance, working memory capacity and cortical thickness. Therefore, we tested a combined model including all predictors to account for neural activity in the three ROIs. The model-building procedure consisted of multiple steps. We started with the best fitting age-model (linear, quadratic or cubic) for the ROI determined in the previous analysis. To test whether other measurements explained additional variance above age, we first added task performance as a second predictor to investigate whether performance explained additional variance above age. If this was the case, performance was included in the next step and if this was not the case, we continued to the next step with only the linear, quadratic or cubic age term. The next step was adding working memory capacity to the model, followed by cortical thickness of the ROI as a final step. AIC and log-likelihood values were used to test whether model fit improved by adding the predictor. Changing the order of adding performance, working memory and cortical thickness did not change the overall results.

Reliability

For visualization purposes, the whole-brain results for the Learning > Application contrast for TP1 and TP2 across all participants are displayed in Figure 2a and Table 1. The results showed that a comparable network was recruited on TP1 and TP2, including bilateral DLPFC, pre-SMA/ACC, and bilateral SPC.

Table 1: MNI coordinates of local maxima activated for the contrast Learning > Application averaged across the two measurements, across all participants, FWE-corrected at $p < .05$.

Area of activation	x	y	z	voxels	T
R inferior parietal lobule	45	-43	49	22000	33.64
R supplementary motor area	3	17	49	s.c.	31.87
R middle frontal gyrus	45	29	34	s.c.	31.79
R superior frontal gyrus	27	2	58	s.c.	30.33
R angular gyrus	33	-58	49	s.c.	28.34
Supplementary motor area	0	11	55	s.c.	27.65
L supplementary motor area	-3	8	58	s.c.	27.17
R insula lobe	33	23	1	s.c.	26.91
R precuneus	6	-70	55	s.c.	26.60
R precuneus	9	-73	58	s.c.	26.13
L middle frontal gyrus	-45	23	37	s.c.	25.93
L insula lobe	-30	20	4	s.c.	25.74
L middle frontal gyrus	-48	26	34	s.c.	25.68
L inferior parietal lobule	-48	-43	46	s.c.	25.46
L inferior parietal lobule	-42	-52	55	s.c.	25.04
R middle frontal gyrus	33	53	13	s.c.	25.01
R posterior cingulate cortex	3	-31	25	26	7.21
Cerebellar vermis	0	-52	-8	13	5.68

Abbreviations: L = Left; R = Right; s.c. = same cluster

For all predictors (ROI activity and cortical thickness, task performance and working memory capacity), we also assessed reliability in ROIs from TP1 to TP2.

We calculated intra-class correlation coefficients (ICCs) on the mean signal of each ROI. We used a two-way mixed model with absolute agreement and we reported the average measure. A value of 0 indicates no relation between the first and second time point and a value of 1 indicates perfect agreement. Interpretation of ICC values for reliability was guided by Cichetti (2001): values $< .4$ were interpreted as poor; values $.41-.59$ were interpreted as fair, values $.60-.74$ were interpreted as good, and values $> .75$ were interpreted as excellent.

As can be seen in Figure 3, most ICC values were in the ‘fair to good’ range. For a visual comparison of ICCs across different age groups, Figure 3 also shows ICC values for 3 age groups (8-12 years, 13-16 years, 17-25 years). For DLPFC, reliability was lowest (in the poor range) in the oldest age groups (suggesting change over time), and was fair for the children and adolescents. In contrast, for SPC and pre-SMA/ACC, reliability was lowest in the child group (suggesting change over time), and was fair to good for adolescents and adults.

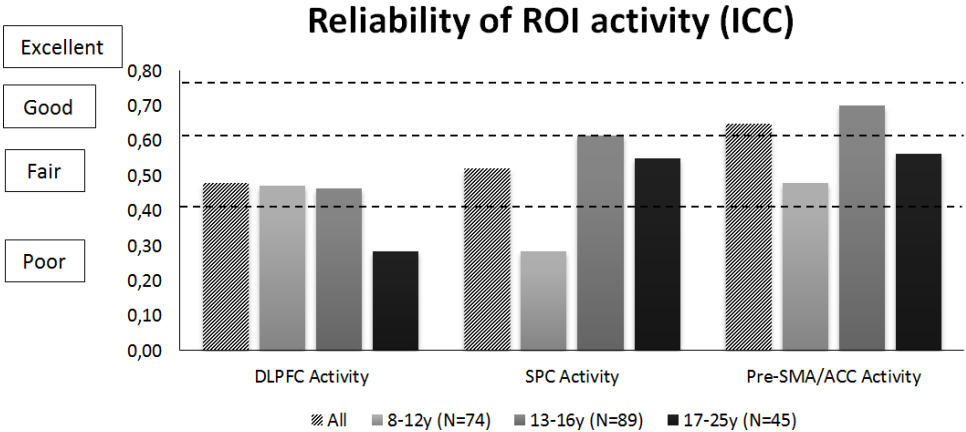


Figure 3: Intra-Class Correlations (ICC) values for each ROI per age group. The labels ‘poor’, ‘fair’, ‘good’ and ‘excellent’ are based on Cicchetti (2001).

Results

Developmental trajectories

For descriptive purposes, correlations between measures are presented in Table 2. Next, we started the model building procedure. As a first step, we tested for all measures (neural activity, task performance, working memory capacity and cortical thickness) whether a linear, quadratic or cubic age pattern best described developmental change. AIC values and a log-likelihood test were used to test which model best fit the data. AIC values for the base model (without age), linear, quadratic and cubic age model for each measure are listed in Table 3.

For neural activity (i.e., the Learning > Application contrast), we observed distinct developmental changes across the different ROIs. For DLPFC, the relationship between age and neural activity was best described by both a linear and a quadratic term for age, i.e. activity increased with age, and leveled off in late adolescence and young adulthood (see Figure 4). For SPC, there was a quadratic but not a linear effect of age, which indicated that activity increased until adolescence and then decreased into adulthood. In the pre-SMA/ACC, neural activity was best described by a linear effect of age that showed increasing neural activity with increasing age. In all ROIs, there was significant individual variability in mean neural activation as indicated by a random effect of the intercept. See Table 3 for the model comparison values and Figure 4 for the predicted and actual data.

Table 2: Correlations between all variables at both time points.

Time point 1	DLPFC	SPC	SMA	DLPFC CT	SPC CT	SMA CT	WM	Perf	Age TP1
DLPFC	1								
SPC	.463**	1							
SMA	.489**	.574**	1						
DLPFC CT	-.192**	n.s.	n.s.	1					
SPC CT	n.s.	n.s.	n.s.	.390**	1				
SMA CT	n.s.	n.s.	n.s.	.423**	.332**	1			
WM	.203**	n.s.	n.s.	n.s.	n.s.	-.144*	1		
Perf	.343**	.222**	.155*	-.140*	n.s.	n.s.	.437**	1	
Age TP1	.410**	.205**	.210**	-.382**	-.162*	-.307**	.340**	.444**	1
Time point 2	DLPFC	SPC	SMA	DLPFC CT	SPC CT	SMA CT	WM	Perf	Age TP2
DLPFC	1								
SPC	.326**	1							
SMA	.400**	.389**	1						
DLPFC CT	-.146*	n.s.	-.170*	1					
SPC CT	n.s.	n.s.	n.s.	.383**	1				
SMA CT	n.s.	n.s.	n.s.	.491**	.341**	1			
WM	.165*	n.s.	n.s.	n.s.	n.s.	n.s.	1		
Perf	.227**	n.s.	.159*	n.s.	n.s.	n.s.	.285**	1	
Age TP2	.209**	n.s.	.296**	-.380**	n.s.	-.307**	.311**	.167*	1

***: Significant at $p < .01$, **: Significant at $p < .05$. Abbreviations: CT = cortical thickness, WM = working memory, Perf = performance, n.s. = not significant.**

Table 3: AIC and loglikelihood p -values for the base model, and linear, quadratic and cubic age models for brain activity, cortical thickness, task performance and working memory capacity.

	<u>Base</u>	<u>Linear</u>		<u>Quadratic</u>		<u>Cubic</u>	
	AIC	AIC	p	AIC	p	AIC	p
DLPFC activity	1562.10	1537.33	<.001	1533.26	.014	1535.09	.675
Pre-SMA/ACC activity	1872.18	1851.03	<.001	1851.70	.249	1853.45	.618
SPC activity	1688.66	1690.18	.486	1683.57	.003	1685.30	.600
DLPFC CT	-425.54	-553.83	<.001	-553.17	.248	-553.715	.110
Pre-SMA/ACC CT	-336.29	-421.659	<.001	-421.60	.164	-430.81	<.001
SPC CT	-447.86	-478.99	<.001	-488.05	<.001	-486.25	.656
Performance	2442.43	2406.12	<.001	2387.45	<.001	2386.47	.084
Working Memory	-485.96	-529.56	<.001	-546.92	<.001	-548.52	.057

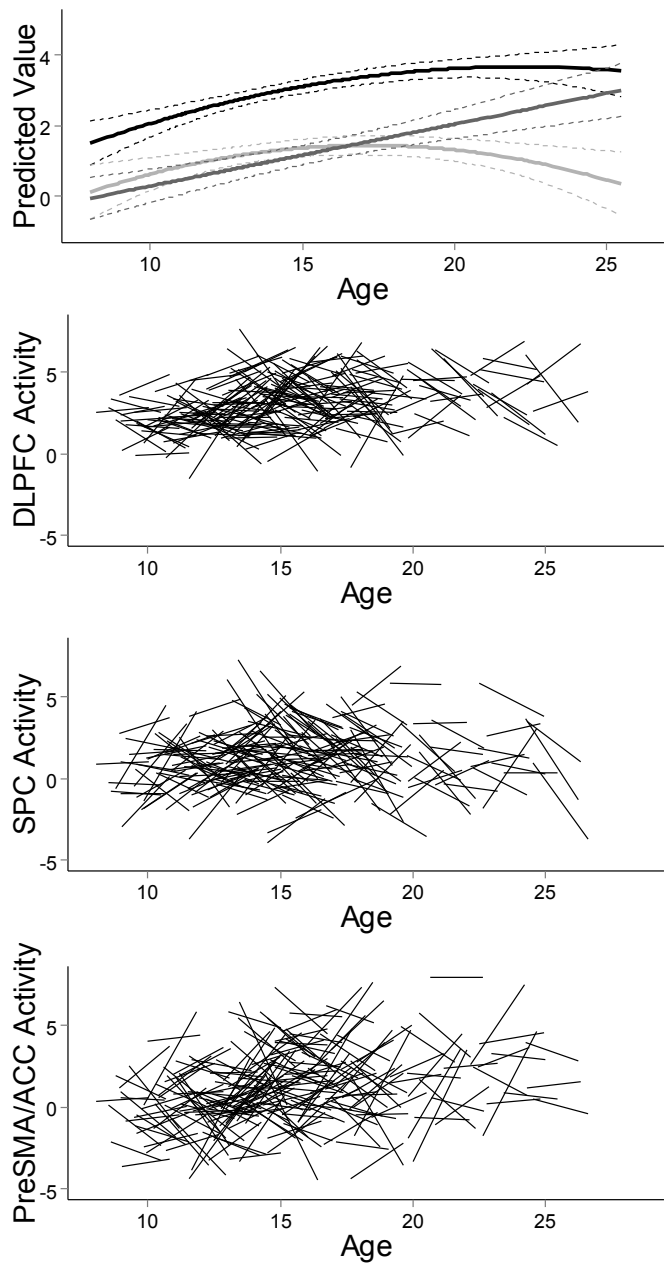


Figure 4a: Predicted data (presented in the upper graphs) and raw data for neural activity in DLPFC, SPC and pre-SMA/ACC. Each line represents one individual at two time points.

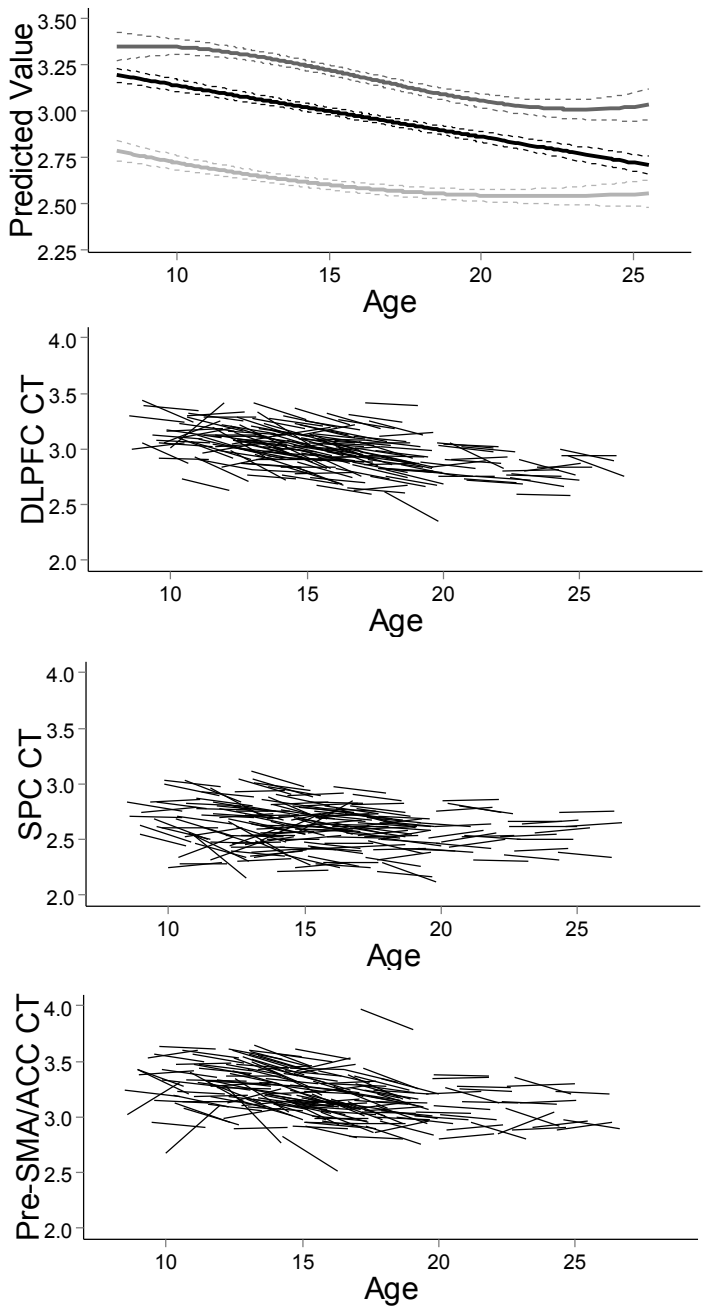


Figure 4b: Predicted data (presented in the upper graphs) and raw data for cortical thickness (CT) in DLPFC, SPC and pre-SMA/ACC. Each line represents one individual at two time points.

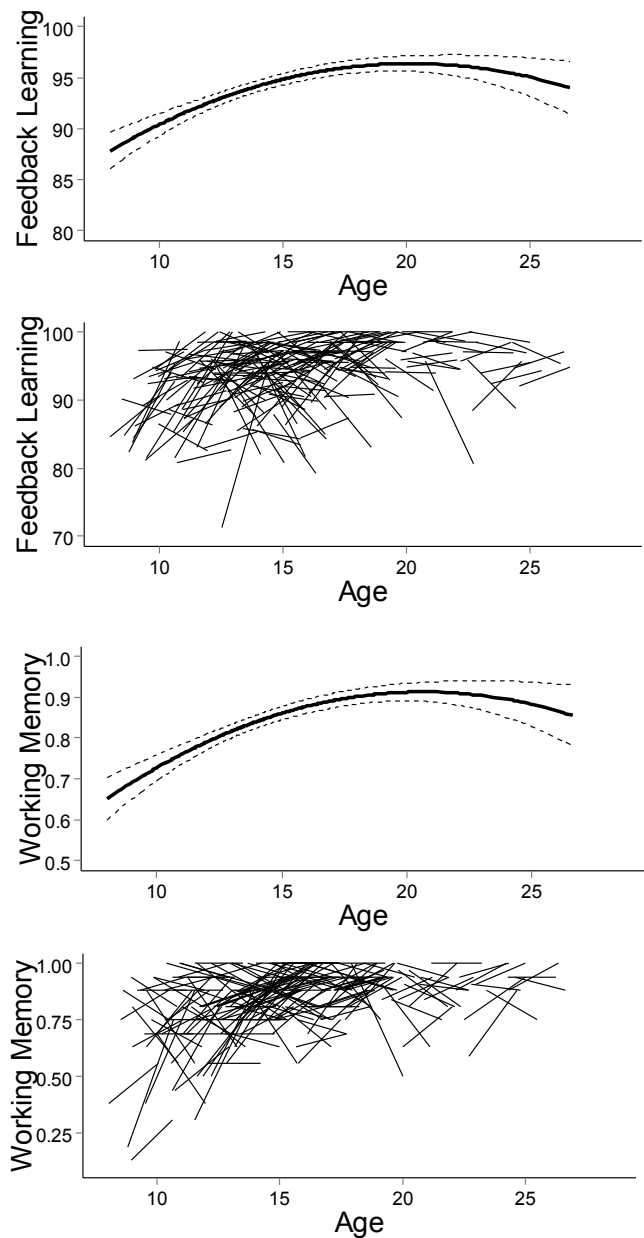


Figure 5: Predicted data and raw data for feedback learning performance and working memory capacity. Feedback learning performance was defined as the percentage of trials during the learning phase where (positive or negative) feedback was successfully used on a subsequent trial.

For task performance, there was a linear and quadratic effect of age on task performance. Note that although the AIC value was slightly lower for a cubic model than for a quadratic model, this difference was not significant according to a log-likelihood test ($p = .084$). This indicated that task performance improved with age, and then leveled off for older participants (see Table 3 and Figure 5 for the raw data and predicted data).

The age models for working memory indicated that a combined linear and quadratic model described the best fit. Note that the AIC value was slightly lower for the cubic age model, but the accompanying log-likelihood test showed no significant improvement ($p = 0.58$) (see Table 3; Figure 5). This indicated that working memory performance improved with age, and then leveled off for older participants.

For cortical thickness of the ROIs, we found different developmental patterns depending on the region. For DLPFC, a linear model with decreasing cortical thickness with age best described the data. In contrast, in SPC the relationship between age and cortical thickness was best described by a quadratic model with a significant linear and quadratic term for age. That is, cortical thickness decreased with age and stabilized in late adolescence/young adulthood (see Figure 4). For pre-SMA/ACC, a model with both a linear and cubic (but not quadratic) effect of age best described the data, i.e., cortical thickness showed a relatively stable pattern in young adolescents, decreased steeply in adolescence and stabilized in late adolescence/early adulthood (see Figure 4). All results are described in Table 3 and visualized in Figure 4.

Explaining developmental change with age, performance, working memory and cortical thickness

To investigate which factors additionally drive these developmental changes in neural activity, we tested the contributions of task performance, working memory capacity, and cortical thickness in addition to age in a hierarchical mixed regression model. Neural activity (i.e., the Learning > Application contrast) was the dependent variable (for each ROI separately) and the predictors task performance, working memory and cortical thickness were added in a consecutive order above age. The starting model was the model with the best fitting age shape (linear, quadratic or cubic). Next, task performance, working memory and cortical thickness were added in hierarchical steps (see Methods section). For DLPFC, the model that best explained neural activity was a model including a linear (but no longer quadratic) age effect and a positive effect of task performance, i.e. better performance predicted increased activity. Cortical thickness and working memory did not explain additional variance over and above age and task performance. For SPC, both age and task performance explained a significant amount of variance in neural activity, such that in addition to a quadratic age effect, better task performance was associated with increased activity. Similar to the DLPFC, working memory capacity and cortical thickness did not contribute to the model over and above age and task performance. A different pattern was found for pre-SMA/ACC. Here, we observed a significant positive linear effect for age and cortical thickness,

but not for task performance and working memory. This model indicated a positive relation between cortical thickness and neural activity, such that increased activity was associated with increased cortical thickness. The final model parameters are summarized in Table 4.

Table 4: Model parameters for the best fitting model for DLPFC, SPC and pre-SMA/ACC.

Area	Variance	θ	p	95% CI	
				Lower	Upper
DLPFC					
Random effect					
Intercept	0.69			0.49	0.98
Fixed effects					
Intercept		3.09	<.001	2.92	3.25
Age		5.89	<.001	2.57	9.22
Age ²		-2.17	.18	-5.35	1.00
Performance		0.07	<.001	0.035	0.11
SPC					
Random effect					
Intercept	1.09			0.087	1.36
Fixed effects					
Intercept		1.18	<.001	0.96	1.39
Age		-0.46	.833	-4.74	3.92
Age ²		-4.36	.036	-8.4	-0.31
Performance		0.05	.044	0.0015	0.09
Pre-SMA/ACC					
Random effect					
Intercept	1.51			1.26	1.82
Fixed effects					
Intercept		1.29	<.001	1.01	1.56
Age		0.20	<.001	0.12	0.28
Cortical Thickness		1.48	.042	0.062	2.85

Discussion

The main aim of this study was to examine the developmental trajectory of the frontoparietal network during a feedback-learning task that captures multiple aspects of cognitive control. We tested for different developmental trajectories (linear, quadratic and cubic) in neural activation and for factors contributing to time-related changes in brain activity for cognitive control above age, particularly task performance, working memory capacity and cortical thickness as an index of structural brain development. The results showed that 1) neural activity during a feedback learning task was best characterized by a quadratic age function in DLPFC and SPC, increasing in early adolescence and leveling off/decreasing into adulthood, but a linear pattern for pre-SMA/ACC, i.e. increasing monotonically across adolescence, 2) Task performance explained additional variance above age in DLPFC and SPC, but not pre-SMA/ACC activity, with higher performance predicting increased neural activation, 3) Cortical thickness explained additional variance in pre-SMA/ACC but not in DLPFC and SPC activity, with increased cortical thickness associated with increased neural activation during feedback learning, and 4) There was fair to good reliability for activity in DLPFC, pre-SMA/ACC and SPC across a two-year period for participants aged 8-27 years. These findings are discussed in more detail in the next paragraphs.

Growth trajectories for neural activity, cortical thickness, performance and working memory

Mixed model analyses of the longitudinal data showed that neural activity was described by different patterns in the neural regions of interest. The linearly increasing pattern in pre-SMA/ACC fits with prevailing developmental theories of monotonously increasing cognitive control. However, the quadratic pattern in DLPFC and SPC is more novel and contradicts initial theories of prefrontal cortex maturation, which suggested a linearly protracted developmental pattern until the early twenties. A closer inspection of the studies reported to date shows that no prior large-scale study included our age range of 8-27 years. Instead, prior studies selected age groups, such as adolescents aged 13-17 versus adults aged 25-30 years or other age selections that did not span the whole range of adolescence and early adulthood (Geier, Garver, Terwilliger, & Luna, 2009; Thomason et al., 2009; van den Bos et al., 2009; van Duijvenvoorde et al., 2008; Velanova, Wheeler, & Luna, 2008). Thus, these prior studies may have been underpowered to detect a quadratic pattern, especially because most prior studies used cross-sectional comparisons (Fjell et al., 2010).

Previous studies found contradicting patterns for frontoparietal recruitment, with some showing increases in activation with increasing age and others showing decreases, and both patterns were interpreted as reflecting immaturity in adolescence (Pfeifer & Allen, 2012). Intriguingly, the current findings show peak activity in SPC and to a lesser extent in DLPFC during late adolescence when learning from feedback. For SPC, there was quadratic pattern with a peak in mid-adolescence, whereas for DLPFC, there was a combined linear and quadratic pattern with

peaking and leveling off in late adolescence/early adulthood. Recent research also showed that complex paradigms such as divergent thinking (Kleibeuker, De Dreu, & Crone, 2013) resulted in stronger DLPFC activity in adolescents than in adults. Studies using more basic cognitive control tasks, however, have not reported these peak activations in frontoparietal regions in adolescence (Klingberg et al., 2002; Rubia et al., 2006). Therefore an important direction for future research is to unravel whether, when, and how DLPFC and SPC show peak activity in late adolescence. Recently, it was found that adolescents recruit DLPFC more strongly than adults when financial incentives were offered for performing well (Teslovich et al., 2014), suggesting that adolescents may engage DLPFC more in a context of high motivation. Possibly, the finding of enhanced activity in SPC and DLPFC in late adolescence indicates that this is a time window when the frontoparietal network can be optimally recruited (Crone & Dahl, 2012). This hypothesis should be tested in more detail in future studies using a larger variety of paradigms, and by also including broader age ranges including a larger adult sample than our sample. Nevertheless, we suggest these results seem consistent with a new perspective on nonlinear development of the frontoparietal cortex (see also Casey, 2015).

We also examined the effects of individual differences on time-related changes in neural activity besides age, such as cortical thickness, working memory capacity and performance on the feedback learning task. We first tested the general age patterns for these individual differences, and subsequently related these to changes over time in activation patterns. First, for cortical thickness, all areas showed a decrease with age, which fits with prior studies showing an initial increase in childhood followed by cortical thinning in adolescence (Koolschijn & Crone, 2013; Shaw et al., 2013; Tamnes et al., 2010). In this study the decrease in cortical thickness in DLPFC was best described by a linear pattern. For SPC, a combined linear and quadratic decrease best described the data, i.e. a steep decrease in early adolescence which leveled off towards late adolescence. Cortical thickness decreases in pre-SMA/ACC were best described by a combined linear and cubic function, showing relatively stable levels in early adolescence, decreasing steeply in adolescence, and leveling off in young adulthood. The overall pattern of a developmental decrease is consistent with earlier reports.

Second, both performance on the feedback learning task and working memory capacity showed a quadratic pattern for age, with steadily increasing performance in adolescence which leveled off in adulthood. A close examination of the literature shows that few studies examined cognitive control development across the whole age range of adolescence until adulthood. For example, a large and comprehensive study on the development of executive functions compared children of 7, 11, 15 and 21 years (Huizinga et al., 2006), but not the intermediate ages. Likewise, a study on performance monitoring previously compared performance of 8-9 year-olds, 11-13-year-olds and 18-25-year-old adults, but did not separate between ages within the adult group (van Duijvenvoorde et al., 2008). The strength of the current study is that quadratic changes in task performance and working memory capacity were observed using a longitudinal design.

After testing these general age patterns, we tested whether task performance, working memory capacity and cortical thickness explained additional variance in time-related changes of frontoparietal activity above age. We found that performance explained additional variance above age in DLPFC and SPC but not pre-SMA/ACC activity, with better performance linked to increased activity. This suggests that task performance provides a unique contribution to activity changes that is not captured by age alone. Cortical thickness explained additional variance above age in activity in pre-SMA/ACC but not in DLPFC or SPC. The latter finding is consistent with earlier studies that showed that DLPFC and SPC structure could not explain age differences in neural activity (Haier, Karama, Leyba, & Jung, 2009; Squeglia et al., 2013). The relation between pre-SMA/ACC cortical thickness and neural activity was positive, such that increased cortical thickness (i.e., less mature) is associated with increased activity. This is surprising given that cortical thickness decreased with age. Intuitively, one might expect that pruning of connections during development increases selectivity and effectiveness of synaptic activity and, therefore, leads to stronger activation. Evidence for this line of reasoning, that is, stronger activations are generally found in thinner (i.e. more mature) regions, has been provided by such relationships in frontoparietal regions in children performing linguistic tasks (Lu et al., 2009; Nunez et al., 2011). On the other hand, increases of activation have for instance been reported after training studies (Koelsch, Fritz, Schulze, Alsop, & Schlaug, 2005), in areas that were thicker in musicians than non-musicians (Bermudez, Lerch, Evans, & Zatorre, 2009). Interestingly, our findings are consistent with a study in adult participants by Hegarty et al. (2012) who also found a positive relation between cortical thickness and activity in pre-SMA/ACC, but not in other prefrontal areas.

A limitation of this study is that the participants who were excluded for the cortical thickness analyses due to lower quality data ($n = 28$), were younger on average compared to the included participants, which is unfortunately a common problem in developmental fMRI studies. However, the remaining sample for cortical thickness was still large ($n = 180$; ages 8-27) and this was one of the first large-scale developmental longitudinal studies assessing the contribution of structural maturation to development of brain function. Taken together, our findings build upon prior research showing relationships between brain structure and function, but the underlying mechanisms and models for these relationships are still open for discussion.

Reliability of neural activity across child and adolescent development

Finally, in this study we measured consistency over time of activity in the frontoparietal network within individuals. There is a growing number of longitudinal studies in adults that examined test-retest reliability in the cognitive control network. The results indicate that over periods of two weeks, reliability is good in DLPFC and SPC during an n-back working memory task (Plichta et al., 2012), there is modest to good reliability of DLPFC activity in visual working memory over a period of three months (Zanto, Pa, & Gazzaley, 2014) and there is modest to good reliability of DLPFC and SPC in a feedback monitoring task over a period of three years (Koolschijn et al.,

2011). The question of reliability in developmental populations has not been consistently examined, but a recent study suggested fair reliability in DLPFC, SPC and ACC in 123 participants between ages 9 to 29 in an oculomotor inhibition task (Ordaz et al., 2013). The current findings demonstrate that in a large sample of 208 participants, reliability across two years is fair to good across ages 8-27 years in bilateral DLPFC, SPC and pre-SMA/ACC. The relatively high reliability of cortical regions is consistent with prediction studies that have shown that future academic achievement can be predicted by activity in SPC one year earlier (Emerson & Cantlon, 2014) and two years earlier (Dumontheil & Klingberg, 2012) during working memory tasks, suggesting that activity in DLPFC and SPC is related to future cognitive outcome.

Conclusions and future directions

This study moved beyond prior cross-sectional comparisons by fitting growth curves based on longitudinal data and thereby improved power for detecting developmental change. An interesting direction for future research is to test how neural activity predicts future changes in behavior. A prior study highlighted the important role of the parietal cortex for predicting future behavior such as academic performance (Dumontheil & Klingberg, 2012). However, other studies showed that activity in the frontoparietal network correlated with working memory capacity across sessions, but activity in subcortical areas (basal ganglia and thalamus) predicted future working memory capacity (Darki & Klingberg, 2014; Ullman, Almeida, & Klingberg, 2014).

In future studies it will be of interest to follow individuals in this study for a third time point. With two time points, it was only possible to investigate nonlinear patterns on the group level. In addition, an interesting future direction would be to examine whether and how baseline activity in the frontoparietal network predicts future behavioral outcomes. An important question for future research is to examine not only the spatial, but also the temporal dynamics of cognitive control, such as with event-related potentials. Prior studies have highlighted the feasibility of this approach in young children (Eppinger et al., 2009), and this will allow for the investigation of fast evaluative processes. A specific limitation of this study is that the feedback learning task was relatively easy and performance was near-perfect for older participants. Future studies should investigate whether the results from this feedback-learning paradigm can be replicated using other and more cognitively taxing tasks measuring cognitive control. Finally, we only investigated whether time-related changes in neural activity covary with the factors age, task performance, working memory capacity and cortical thickness. Many other factors may contribute to changes over time in activity, such as increased response inhibition or motivation, which should be investigated in future research.

Taken together, this study accentuates the important role of the emerging frontoparietal network in child and adolescent development. With regard to dual-process models of adolescent development, this study provides evidence against a simple linear development of the frontopa-

rietal network and highlights the need for further large-scale longitudinal studies to test adolescent development more reliably.

