



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 5

Influence of strategy use and age on neural activity for feedback learning



This chapter is based on:

Peters, S., Koolschijn, P.C.M.P., Crone, E.A., van Duijvenvoorde, A.C.K. & Raijmakers, M.E.J. (2014). Strategies influence neural activity for feedback learning across child and adolescent development. *Neuropsychologia*, 62, 365-374.

Abstract

Learning from feedback is an important aspect of executive functioning that shows profound improvements during childhood and adolescence. This is accompanied by neural changes in the feedback learning network, which includes pre-supplementary motor area (pre-SMA)/anterior cingulate cortex (ACC), dorsolateral prefrontal cortex (DLPFC), superior parietal cortex (SPC), and the basal ganglia. However, there can be considerable differences within age ranges in performance that are ascribed to differences in strategy use. This is problematic for traditional approaches of analyzing developmental data, in which age groups are assumed to be homogenous in strategy use. In this study, we used latent variable models to investigate if underlying strategy groups could be detected for a feedback learning task and whether there were differences in neural activation patterns between strategies. In a sample of 268 participants between ages 8 to 25 years, we observed four underlying strategy groups, which cut across age groups and varied in the optimality of executive functioning. These strategy groups also differed in neural activity during learning; especially the most optimal performing group showed more activity in DLPFC, SPC and pre-SMA/ACC compared to the other groups. However, age differences remained an important contributor to neural activation, even when correcting for strategy. These findings contribute to the debate of age versus performance predictors of neural development, and highlight the importance of studying individual differences in strategy use when studying development.

Introduction

An important component of cognitive development is the ability to control and adapt behavior in response to changing environmental demands, also referred to as executive functions (Diamond, 2013; Zelazo, 2006). Executive functions are thought to consist of three core functions: inhibition, working memory and cognitive flexibility (Diamond, 2013). Higher-order executive functions such as reasoning, planning and learning from prior experiences rely upon combinations of these three core functions. The ability to adapt behavior based on prior experiences (i.e. adaptive control) shows a marked improvement during childhood and adolescence (Tamnes, Walhovd, Torstveit, Sells, & Fjell, 2013). For example, in the classic Wisconsin Card Sorting Task (WCST), there is a developmental improvement in flexibly adapting behavior based on positive and negative feedback (Huizinga et al., 2006) and in probabilistic feedback learning tasks there is a developmental improvement in adapting behavior successfully based on informative versus non-informative feedback (Eppinger et al., 2009; Jansen, van Duijvenvoorde, & Huizinga, 2014; van den Bos et al., 2009; Van Duijvenvoorde, Jansen, Griffioen, Van der Molen, & Huizinga, 2013). Despite these convincing developmental patterns, there are large individual differences in adaptive control within age ranges, i.e., not all children and adolescents are equally proficient at learning from positive and negative feedback. Why is it that some children are better at learning compared to their peers? Studying the behavioral and neural mechanisms underlying successful learning is important to advance our understanding of executive control processes and their development.

Most prior studies on the development of feedback learning have focused on performance improvements with age and the accompanying changes in brain activity. Research in adults indicated that during feedback learning, a large brain network is activated, including pre-supplementary motor area (pre-SMA)/anterior cingulate cortex (ACC) (Holroyd et al., 2004; Mars et al., 2005; Monchi et al., 2001; Ullsperger & von Cramon, 2003), (dorso)lateral prefrontal cortex (DLPFC) (Dove et al., 2000; Lie et al., 2006; van Veen et al., 2004; Zanolie et al., 2008), basal ganglia (Monchi et al., 2001; Tricomi et al., 2006), and superior parietal cortex (SPC) (Zanolie et al., 2008). It is thought that a dopamine-initiated alarm signal in pre-SMA/ACC signals that outcomes are worse than expected. Subsequently, the DLPFC is a primary site for implementation of adaptive control (Kerns et al., 2004; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004).

Prior developmental studies have shown that this feedback learning network becomes increasingly activated with age (Crone et al., 2008; Peters, Braams, Raijmakers, Koolschijn, & Crone, 2014; van den Bos et al., 2009; van Duijvenvoorde et al., 2008). However, it is unclear whether these neural changes reflect age differences (i.e., a maturational viewpoint), or whether they are related more to performance differences rather than age (Andersen et al., 2014; Jolles & Crone, 2012; Koolschijn et al., 2011).

Effects of performance versus age are only scarcely investigated in developmental feedback learning studies. Moreover, most studies have assumed that performance differences are continuous, implying that all participants within an age group perform the task using the same strategy. However, performance is not constant within age groups: Some children perform at levels similar to adults, whereas others never seem to reach the highest performing levels. It is possible that these individual differences in performance can be described by differences in strategy use. Such differences in performance and strategy use within age groups pose a considerable problem for traditional ways of analyzing developmental data, because these are based on the assumption of homogenous strategy use within age groups.

A robust approach for analyzing individual differences are categorical latent variable models, which allow for detection of different strategies based on individuals' responses across trials. Such techniques have been applied by a number of studies that distinguished distinct learning strategies within age groups (Andersen et al., 2014; Raijmakers, Dolan, & Molenaar, 2001; Schmittmann et al., 2012, 2006; Speekenbrink, Lagnado, Wilkinson, Jahanshahi, & Shanks, 2010). For instance, Schmittmann et al. (2006) showed that two distinct learning strategies (resulting in relatively fast or slow learning) could be distinguished in a category-learning task. The fast and slow strategy groups both employed a learning strategy based on hypothesis-testing (as opposed to incremental, associative learning), but participants in the slow group were less efficient in their hypothesis testing compared to the fast group. This difference in efficiency was categorical. That is, with age, children were increasingly likely to belong to the faster strategy group; they were not simply less efficient in employing the same strategy. In the current study, we applied these methods to a feedback learning task and investigated whether distinct learning strategies were also observable at the neural level.

In the current paradigm, we built on prior studies on the development of feedback learning, such as a rule switch task used in Crone et al. (2008) and a rule search and application task used by Van Duijvenvoorde et al. (2008), and constructed a paradigm in which correct responses could be inferred through a process of hypothesis-testing. Additionally, different deductive reasoning steps could be applied to use a more efficient hypothesis testing strategy. This made the task suitable for differentiating between categorically different strategies, rather than simply assessing performance differences within one strategy. We asked 268 participants ranging from 8 to 25 years to sort stimuli in one of three locations by using positive and negative feedback. An efficient way of solving this task was to not only focus on feedback for the current stimulus but also to remember the locations for the other two stimuli. We recorded trial-by-trial data on learning efficiency and analyzed this with latent variable modeling approaches (Markov models and finite mixture models), to investigate if latent strategy groups could be detected (van der Maas & Straatemeier, 2008). As a further addition to prior research, we investigated if underlying strategy groups could be distinguished at the neural level (see also Andersen et al., 2014). We hypothesized that age differences in neural activity for feedback learning are largely attributable

to differences in strategy use. Thus, we tested whether age differences in neural activity were influenced by strategy use, or if there was also neural activity related to maturational processes per se, independent of strategy use. The main developmental effects have previously been reported in Peters et al., 2014. This dataset presents a unique opportunity for analyzing strategy-related versus age-related neural changes in feedback learning given the large-sample size across a broad developmental range.

Methods

Participants

The sample included 268 participants (138 female) between 8.01 and 25.95 years old ($M = 14.22$, $SD = 3.63$), who were recruited through local schools and advertisements. See Table 1 for the number of participants per age and per sex.

Table 1: Number of participants per age and sex.

Age	N Female	N Male	N Total
8 years	6	4	10
9 years	14	5	19
10 years	11	12	23
11 years	13	14	27
12 years	19	11	30
13 years	16	20	36
14 years	10	17	27
15 years	10	11	21
16 years	11	9	20
17 years	12	11	23
18-25 years	16	16	32
N Total	138	130	268

Adult participants (18-25 years) were grouped together. A chi square test indicated that the proportion of males to females was similar across age groups ($\chi^2(10) = 9.20$, $p = .514$). IQ scores were estimated with two subtests of the WAIS-III or WISC-III (Similarities and Block Design). Estimated IQ scores ranged from 80 to 143 ($M = 110.25$, $SD = 10.62$) and showed no correlation with age ($r = -.09$, $p = .155$). None of the participants reported a history of neurological or psychiatric disorders or current use of psychotropic medication. All anatomical MRI scans were reviewed and cleared by a radiologist. The study was approved by the Institutional Review Board at the Leiden University Medical Center and all participants (or participant's parents for minors) provided

written informed consent. Adults received payment for participation, and children and their parents received presents and a fixed payment for travel reimbursement.

Exclusion criteria

Twenty-five participants were excluded (not included in Table 1) from further analyses after participation for the following reasons: Nineteen participants were excluded because movement in the MRI scanner exceeded 3.0 mm in any direction, three participants were excluded because of technical problems and three participants were excluded because they were outliers (more than three times the interquartile range) on the total percentage of positive feedback, indicating they did not perform the feedback learning task adequately.

Feedback Learning Task

Participants performed a feedback learning task in the MRI scanner (see also Peters et al., 2014). On each trial, they saw three empty squares, under which one of three different stimuli was presented (see Figure 1). We explained to the participants that each stimulus belonged in one of the three squares and they had to find the correct location for all stimuli by using performance feedback. Performance feedback was a plus-sign for positive feedback and a minus-sign for negative feedback. After either 12 trials, or when the participant correctly applied the correct location twice in total for each stimulus, the sequence ended and a new sequence was presented with three new stimuli. There were 15 sequences in total, resulting in a maximum of 180 trials. Stimuli were presented in a pseudorandom order, with a maximum of two identical stimuli in a row. Before the MRI session, all participants practiced three sequences. During the MRI session the task was divided into two runs of eight and seven sequences. Each trial started with a 500 ms fixation cross. Consecutively, stimuli were presented for 2500 ms, during which time window the response had to be given. Participants saw the words “Too Late” if they did not respond within this time window, after which the sequence continued. After the response, performance feedback was presented for 1000 ms. Inter-trial intervals were jittered to optimize the timing for fMRI based on OptSeq (Dale, 1999) with intervals between 0 and 6 seconds.

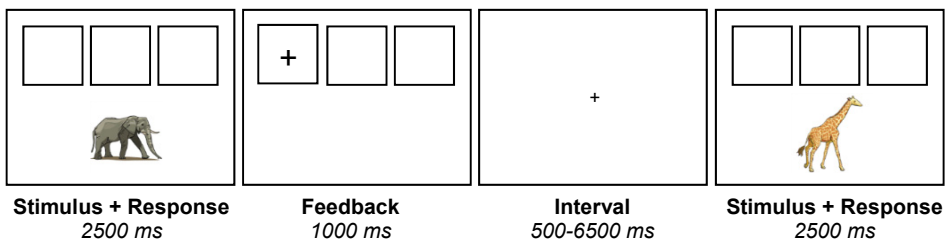


Figure 1: Display of task sequence.

Modeling strategies

To model latent behavioral strategies for the task we first recoded the trial-by-trial accuracy data into new trial-by-trial data, which categorizes different response types that revealed reasoning level. Each trial was scored as one of these four response types, in ascending order of the complexity of reasoning involved:

(a) *Mistake*: repeating a previously made error or making an error after an earlier correct response for the same stimulus. This can be seen as a short-term memory (STM) error (Diamond, 2013), where the subject has forgotten the correct answer it has seen before; no active working memory (WM) calculations were necessary to avoid a mistake. (b) *Inefficient*: when the location of one picture was known, the participant failed to deduce that this location could not be the correct location for another picture. This could be described as a working memory error. In this trial the subject could have avoided an error by means of simple WM calculations. (c) *Suboptimal*: when participants received negative feedback for a stimulus in one location, the optimal decision is to place another stimulus in that location, to ensure a 50 percent chance of being correct (instead of 33 percent). Failing to use this strategy is suboptimal in reasoning, i.e., the subject did not optimize the probabilities of a correct choice. The avoidance of suboptimal trials involves complex executive functions, such as planning and reasoning (Diamond, 2013). (d) *Optimal*: all other cases, i.e. choosing the best possible option given the information that is acquired from previous trials.

The resulting trial-by-trial data on response types was used to distinguish individuals with distinct learning strategies. To detect latent strategies in trial-by-trial data, we used two types of categorical latent variables models. The first type is a static, finite mixture distribution model. Finite mixture distribution models group individuals per learning strategy (based on their response pattern), which is assumed to be constant across the task. The second type is a dynamic Markov model that defines changes in learning strategies across trials (Rabiner, 1989; Visser, 2011). Simulation studies show that such latent variable models are robust statistical techniques, which are necessary to make a reliable decision about the number of strategies and about the nature of the strategies, compared to more ad hoc methods for distinguishing between different learning strategies (van der Maas & Straatemeier, 2008). For fitting models to the data, we calculated maximum likelihood estimates of the parameters in the model by using the statistical R-package *depmixS4* (Visser & Speekenbrink, 2010). To determine the most parsimonious, best fitting model to the data, i.e., the optimal model, the Bayesian Information Criterion (BIC) (Schwarz, 1978) was used.

First, we used static, finite mixture distribution models to determine whether performance differences could best be described by either a continuous variation in performance (with age as a covariate) or by a number of categorically different strategies. The former is modeled by including age as a covariate on the response probabilities. According to this model, each individual responds with specific probabilities for the four response types (optimal, inefficient, mistake, suboptimal) and this response pattern depends on age, but without the presence of any latent

strategies. Next, we fitted finite mixture distribution models to the data with a varying number of groups. Comparing these models by means of BIC will show if a model with a certain number of categorically different strategies is better than one with continuous performance differences. Second, we applied a dynamic Markov model to the data. Markov models define changes in learning strategies across trials (Rabiner, 1989; Visser, 2011). We thus extended the mixture of static models (first class of models) to a mixture of (dynamic) Markov models to allow for changes during the learning process. The first Markov models include different states within a learning event. These models could describe, for example, subgroups of participants switching between an inefficient towards an efficient state within a sequence (i.e., different phases during one learning event). The second type of Markov models we tested included a continuous change of strategies across trials within a sequence (i.e., for one or more latent strategies, response probabilities continuously vary with trial number).

Model-comparisons by means of BIC indicated that models with the most parsimonious strategies, corresponding to the firstly described static finite mixture distribution models were optimal for this data compared to the dynamic Markov models. In the results section, we will therefore focus on the static finite mixture distribution models and discuss the number of latent strategies (one to five) that is optimal to describe the data. When the results indicated that the optimal model for the data consisted of multiple strategies, we assigned individuals to the strategy that they were most likely applying based on the posterior probabilities of the data given the model (Visser, 2011). Subsequently, this assignment of individuals was used in the fMRI analysis.

fMRI analyses: Learning and Application

For the fMRI analysis, we used a contrast that reveals brain areas with sensitivity to feedback with learning value. That is, we aimed to find areas that respond more to feedback that provides new information, compared to feedback that provides information that is already known. In our opinion, distinguishing between ‘useful’ and ‘less useful’ feedback is one of the key aspects of feedback learning, which is why we expected that this contrast is related to individual differences in performance and strategy use. To identify ‘useful’ and ‘less useful’ feedback, we distinguished between a learning phase and an application phase for each stimulus. The learning phase was defined as those trials in which participants had not yet responded with the correct location for the stimulus, and were thus still trying to find the correct solution. We only included trials that actually resulted in learning ($M = 96.35\%$, $SD = 3.03\%$ of all trials). For learning from positive feedback (Positive Learning), this meant choosing the same location on a next trial for that same stimulus, and for learning from negative feedback (Negative Learning), that meant *not* choosing the same location on a next trial. The application phase was defined as those trials in which a stimulus was sorted correctly on a preceding trial, and which continued to be sorted correctly. All analyses were based on the contrast Learning (Positive Learning & Negative Learning) > Application, in which Positive Learning and Negative Learning were combined and compared to Appli-

cation. In doing this we followed a similar approach to a prior study focusing on feedback differing in informative value for learning (Eliassen et al., 2012). One potential confound is that effects are due to negative feedback rather than feedback which signals learning. However, whole-brain results were highly similar when the Learning (Positive Learning & Negative Learning) > Application and the contrast Positive Learning > Application are compared (see Figure 2). We additionally tested if including negative feedback in the contrast did not lead to additional activation when compared to Positive Learning > Application. Thus we tested with an exclusive mask if Learning (Positive Learning & Negative Learning) > Application showed any significant activation that was not in the contrast Positive Learning > Application. This analysis did not result in significant remaining activations. From now on, the contrast Learning (Positive Learning & Negative Learning) > Application will be abbreviated to Learning > Application.

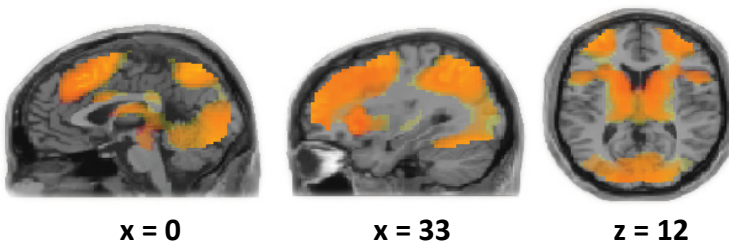


Figure 2: Overlay showing activity for Positive Learning > Application (yellow), Learning (Positive Learning & Negative Learning) > Application (red) and their overlap (orange) ($N = 268$). Both contrasts were FWE-corrected, $p < .05$, > 10 contiguous voxels.

Data Acquisition

MRI scans were acquired with a standard whole-head coil on a Philips 3.0 Tesla MRI scanner. Functional scans were acquired during two runs with T2*-weighted echo-planar imaging (EPI). The first two volumes were discarded to allow for equilibration of T1 saturation effects. Volumes covered the whole brain ($TR = 2.2$ s, $TE = 30$ ms, sequential acquisition, 38 slices, slice thickness = 2.75 mm, Field of View (FOV) = $220 \times 220 \times 114.68$ mm). A high-resolution 3D T1-FFE scan for anatomical reference was obtained after the experimental tasks ($TR = 9.76$ ms, $TE = 4.59$ ms, 140 slices, voxel size = 0.875 mm, FOV = $224 \times 177 \times 168$ mm). The experimental task was projected on a screen that was viewed through a mirror. Before the MRI scan, participants were accustomed to the MRI environment and sounds with a mock scanner.

FMRI Data Analysis

All data were analyzed with SPM8 (Wellcome Department of Cognitive Neurology, London). Images were corrected for slice timing acquisition and rigid body motion. Structural and functional volumes were spatially normalized to T1 templates. The normalization algorithm used a 12-parameter affine transform together with a nonlinear transformation involving cosine basis functions and resampled the volumes to 3 mm cubic voxels. 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 8mm FWHM isotropic Gaussian kernel. The fMRI time series data were modeled by a series of events convolved with a canonical hemodynamic response function. The modeled events were “Positive Learning”, “Negative Learning”, and “Application”, which were time-locked with 0-duration to the moment of feedback. All other trials (e.g., trials that did not result in learning or too-late trials) were modeled as events of no interest. The trials were used as covariates in a general linear model; along with a basic 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 resulting contrast images, computed on a subject-by-subject basis, were submitted to group analyses. All fMRI analyses were initially calculated with a stringent FWE-corrected threshold at $p < .05$, with at least 10 contiguous voxels. In order to examine whole brain effects in more detail, region-of-interest (ROI) analyses were performed with the Marsbar toolbox in SPM8 (Brett et al. 2002).

Results

Modeling results

We fitted finite mixture distribution models to the trial-by-trial data on response types (see Methods section) with a different number of groups (one to five subgroups). In addition, we fitted a 1-state model (i.e., no underlying strategy groups) to the data that assumes continuous individual variation in behavior related to age. By comparing the BIC between models we selected the optimal model to describe the data. Table 2 shows the fit statistics of the different models. First, it appeared that a continuous variation of response probabilities with age was not the optimal way to describe the data. That is, the 1-state age model does not have the lowest BIC. The optimal model assumed the presence of four latent groups (4-states model). We did not test more than five groups because the BIC increased with five groups compared to four groups. Table 3 shows the number of participants per group and how the four groups are defined by percentages of response types.

As can be seen in Table 3, the four strategy groups differed in the number of trials per trial type. We described these four groups as follows: 1) Low strategy group: Relatively many mistakes (no reasoning required, errors in STM or motor response), inefficient responses (not

taking into account feedback for other stimuli: WM errors) and suboptimal responses (not maximizing chance of correctness). 2) Medium-Suboptimal group: Fewer mistakes and inefficient decisions than the Low group, but still relatively many suboptimal decisions. 3) Medium group: Comparable to the Medium-suboptimal group, but fewer suboptimal decisions. 4) High strategy group: Almost all trials are optimal in terms of strategy use.

Table 2: Fit statistics of the static latent variable models with different number of latent strategies.

Model	Loglike	df	BIC
1 state	-15106.51	3	30244.56
1 state age	-14898.26	6	29859.60
2 states	-14605.19	7	29283.97
3 states	-14496.1	11	29107.84
4 states*	-14461.58	15	29080.87
5 states	-14461.32	19	29122.39

Note: The models are explained in the methods section. A 1-state model assumes one strategy; a 2-state model two strategies, etc. The 1-state age model assumes continuous variation between individuals that is related to age. Loglike is the loglikelihood of the fitted model; df = degrees of freedom, which is in this case the number of freely estimated parameters; BIC = Bayesian Information Criterion.

* The model with the lowest BIC is the optimal model.

Table 3: The number of participants per group and percentage of trials per response type.

Strategy group	N	Mistake	Inefficient	Suboptimal	Optimal
Low	33	7.04 %	8.14 %	7.52 %	77.30 %
Medium-Suboptimal	46	2.12 %	3.58 %	6.77 %	87.53 %
Medium	76	2.91 %	4.51 %	2.22 %	90.36 %
High	113	0.74 %	1.45 %	1.04 %	96.77 %

A repeated-measures analyses with the frequencies of the four response types as the within-subjects variables, and strategy group as between-subjects variable yielded a significant interaction between response-type and strategy group ($F(9, 792) = 109.94, p < .001$), which is to be expected beforehand from the mixture distribution analysis. Further follow-up one-way ANOVA tests indicated that there were significant between-group differences in the frequency of optimal trials ($F(3, 264) = 50.84, p < .001$, with Bonferroni post-hoc tests indicating that all groups differed from each other (all $ps < .002$); except the Medium group and the High group); the frequency of mistakes ($F(3, 264) = 194.15, p < .001$, post-hoc tests showed all groups differed from each other (all $ps < .017$)); the frequency of inefficient responses ($F(3, 264) = 189.44, p < .001$, post-hoc tests indicated that all groups differed (all $ps < .015$)); and finally, the frequency of suboptimal responses ($F(3, 264) = 276.56, p < .001$, with post-hoc tests showing that all groups differed from each other (all $ps < .005$)). These results provide evidence that the High strategy group is most efficient in the

hypothesis-testing process because almost no reasoning mistakes were made. In the next paragraphs, we will describe how the four strategy groups are divided over age groups, and subsequently how the groups differed in brain activity during the feedback learning task.

Division of age and performance over the different strategy groups

In Figure 3, the division of strategy groups over discrete age groups is displayed. A chi-square test indicated that age groups differed across the strategy groups ($\chi^2(30) = 82.80, p < .001$). This indicates that age is an important factor contributing to strategy use. However, even in the two extreme groups (Low and High strategy), children and adults of different ages were distributed over these strategy groups, such that, for instance, some young children (from 9 years onwards) were present in the High strategy group and one adult was present in the Low strategy group. Note that IQ did not differ across strategy group as indicated by a one-way ANOVA ($F(3, 262) = 1.82, p = .145$). The division across strategy groups was similar for males and females ($\chi^2(3) = 0.25, p = .969$).

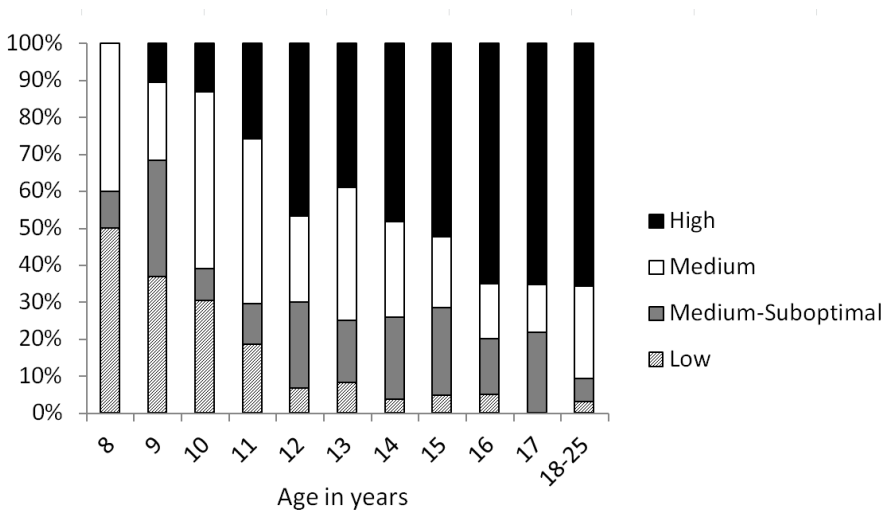


Figure 3: Division of strategy groups (in percentages) over age groups.

FMRI analysis

The percentage of trials for Learning and Application out of the total number of trials per strategy group are displayed in Table 4. Note that the total number of trials could differ per participant due to the nature of the task. ‘Other trials’ are trials which did not result in learning, incorrect

applications or trials where the participant responded too late. These trials were not used for further fMRI analyses.

Table 4: Percentage of trials for Learning and Application (Mean (SD))

Strategy Group	Positive Learning	Negative Learning	Application	Other trials
Low	24.86 (2.66)	25.36 (3.93)	35.65 (3.77)	14.13 (4.81)
Medium-SO	29.98 (2.18)	21.33 (2.63)	42.82 (2.35)	5.87 (2.37)
Medium	29.23 (2.47)	19.92 (2.06)	42.93 (2.39)	7.91 (3.01)
High	32.87 (2.02)	17.46 (2.02)	46.85 (1.68)	2.82 (1.91)

To assess potential neural differences between strategy groups, we focused on the contrast Learning > Application. General age and main effects of feedback learning were also described in Peters et al. (2014). Here, we focused on the strategy versus age related differences in neural activation. Our hypothesis was that age effects for the contrast Learning > Application can partly be explained by strategy differences. To investigate this, we first calculated the contrast Learning > Application with age as a positive regressor, to see which areas were more active with increasing age (see Figure 4, Table 5).

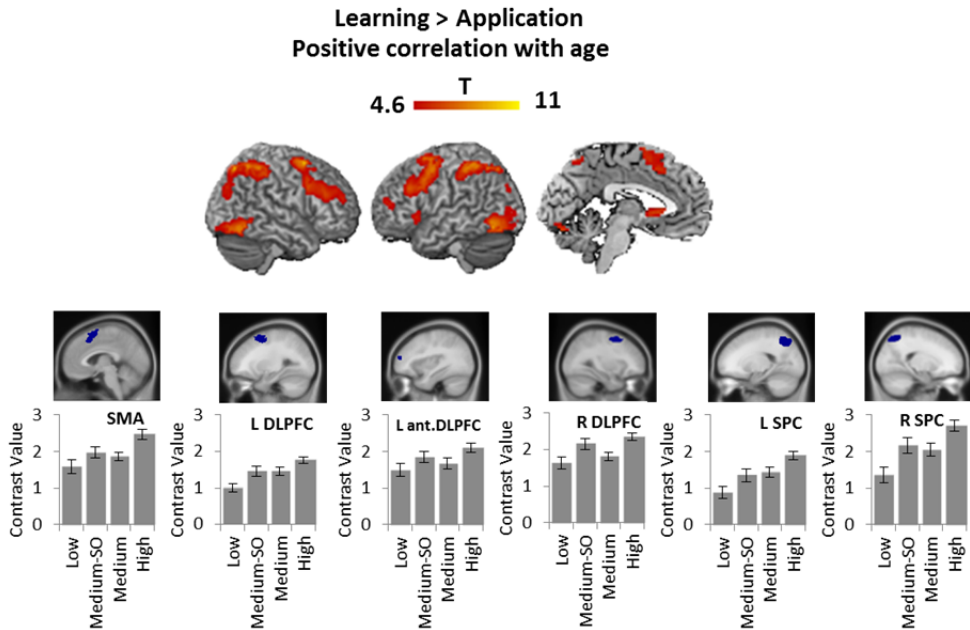


Figure 4: Areas showing activation for the contrast Learning > Application, with age as a positive regressor, FWE corrected at $p < .05$, > 10 contiguous voxels. Differential activity for Learning > Application in ROIs based on this contrast is shown per strategy group (SO = Suboptimal).

Table 5: Areas showing activation for the contrast *Learning > Application*, with age as a positive regressor, FWE corrected at $p < .05$, > 10 contiguous voxels.

Area of activation	x	y	z	voxels	T
R inferior parietal lobule	42	-39	48	2127	10.15
R superior parietal lobule	21	-69	54		9.36
L superior parietal lobule	-27	-63	54		9.07
R middle frontal gyrus	30	9	60	521	9.29
R inferior gyrus	48	12	33		6.57
R inferior gyrus	45	33	24		6.29
L inferior occipital gyrus	-48	-66	-15	407	8.54
L fusiform gyrus	-33	-51	-18		5.54
L middle occipital gyrus	-36	-84	6		5.24
L precentral gyrus	-36	0	60	732	8.24
L precentral gyrus	-48	3	51		7.68
L superior frontal gyrus	-21	12	63		7.37
R inferior temporal gyrus	51	-60	-15	292	8.06
R cerebellum	33	-69	-21		5.75
R inferior occipital gyrus	27	-93	-12		4.96
R caudate nucleus	6	21	0	260	7.00
L caudate nucleus	-9	9	0		6.79
R caudate nucleus	9	9	0		6.53
L cerebellum	-3	-81	-18	52	6.13
L lingual gyrus	-12	-93	-15		4.89
L inferior frontal gyrus	-30	30	0	65	5.58
L inferior frontal gyrus	-45	21	-3		5.42
L middle frontal gyrus	-36	54	15	30	5.50

Note: Abbreviations: L = Left; R = Right

From these functional activations we created ROIs. The resulting ROIs spanned several brain regions, therefore we applied anatomical masks (based on Marsbar Automated Anatomical Labeling) for the key regions implicated in the development of feedback learning (Peters et al., 2014; Crone et al, 2008; Van Duijvenvoorde et al., 2008)): left and right DLPFC (AAL mask: Middle Frontal Gyrus), pre-SMA/ACC (Supplementary Motor Area, left and right combined), and left and right SPC (Superior Parietal Gyrus). Centre-of-mass MNI (x y z) coordinates were: right DLPFC: x = 39, y = 22, z = 41; left DLPFC: x = -35, y = 12, z = 49; left anterior DLPFC: x = -35, y = 52, z = 14 ; pre-SMA/ACC: x = -4, y = 12, z = 58; right SPC: x = 27, y = -62, z = 55; left SPC: x = -24, y = -64, z = 50 (see Figure 4). (Note that the anatomical masking process resulted in two left DLPFC regions, one of which we called: 'left anterior DLPFC').

ROI analyses

To test if strategy groups differed in neural activity within these areas, we created ROIs based on this contrast (see Figure 4). For these ROIs, we first tested whether the four groups differed in neural response; then we tested if strategy group explained variance above age, and finally we tested with mediation analyses whether age effects were mediated by strategy group. Thus we computed difference scores for Learning > Application, which were different for the four strategy groups in all ROIs (all $ps < .005$; see Figure 4). The patterns for each region showed lowest activity for the Low strategy group, higher activity for the Medium and Medium-Suboptimal group, and highest activity for the High strategy group. LSD post-hoc tests for each region separately indicated that for pre-SMA/ACC, all groups showed lower activity than the High group, but there were no other differences (all $ps < .005$). For right SPC, we found that all groups except Medium and Medium-Suboptimal differed from each other (all $ps < .024$), and for left SPC, all groups except Medium and Medium-Suboptimal and Low and Medium-Suboptimal differed from each other (all $ps < .014$). We found that for right DLPFC, the Low and Medium-Suboptimal, Low and High, and Medium and High showed significant differences (all $ps < .017$); for left DLPFC that all groups except Medium and Medium-Suboptimal were different (all $ps < .025$); and for left anterior DLPFC that the Low and Medium group showed less activity than the High group (all $ps < .007$). Note that the general pattern (see Figure 4) is that the Low group shows the least activity and the High group shows the most activity, which for the High group is consistently different from the other groups except in right DLPFC and left anterior DLPFC.

To investigate whether age differences in neural activity can be attributed to differences in strategy use (over and above age), we used hierarchical linear regression analyses with neural activity for Learning > Application (difference score) as dependent variable, age entered as first predictor and strategy group entered as second predictor. We found significant effects of strategy group above age for four of the seven ROIs: pre-SMA/ACC (step 1: $R^2 = .15$; age: $B = .39$, $p < .001$; step 2: $R^2 = .18$; age: $B = .33$, $p < .001$, strategy group: $B = .16$, $p = .007$), left DLPFC (step 1: $R^2 = .22$; age: $B = .47$, $p < .001$; step 2: $R^2 = .24$; age: $B = .43$, $p < .001$, strategy group: $B = .13$, $p = .030$), left SPC (step 1: $R^2 = .23$; age: $B = .48$, $p < .001$; step 2: $R^2 = .25$; age: $B = .43$, $p < .001$, strategy group: $B = .14$, $p = .012$) and right SPC (step 1: $R^2 = .25$; age: $B = .50$, $p < .001$; step 2: $R^2 = .26$; age: $B = .45$, $p < .001$, strategy group: $B = .12$, $p = .033$). To summarize, these results indicate that strategy explained additional variance in neural activity above age in pre-SMA/ACC, bilateral SPC and left DLPFC.

Mediation analyses

For the four ROIs that showed a significant contribution of strategy group above age, we performed mediation analyses with the R package for causal mediation analysis (Imai, Keele, & Tingley, 2010) to investigate the relation between age and strategy use in explaining variance in brain activity. We applied the analysis on ROI activity with age (continuous) as direct predictor and strategy group (nominal variable) as mediator variable (see Figure 5).

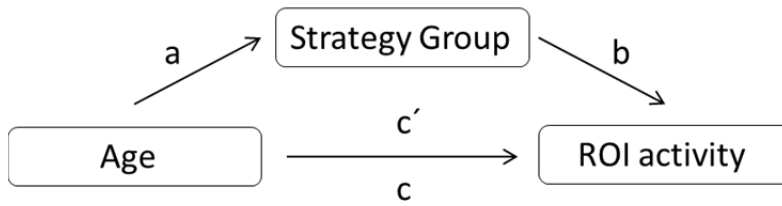


Figure 5: Mediation analysis paths for strategy group as a mediator between age and ROI activity.

With causal mediation analysis, a mediation effect is present if 1) age predicts strategy group (path a), 2) strategy group predicts ROI activity if age is simultaneously entered as a predictor (path b), 3) if age predicts ROI activity (path c), and 4) if strategy group is entered simultaneously as a predictor, the effect of age on ROI activity decreases (path c') (Preacher & Hayes, 2008). We first tested for collinearity problems between age and strategy group, which are indicated by a Variance Inflation Factor greater than 10 (Myers, 1990), a tolerance value less than 0.1 (Menard, 1995) and condition indices greater than 10 (Belsley, David, Kuh & Welsch 1980). For this data, we found no indication for collinearity problems ($VIF = 1.14$, tolerance = 0.86, condition indices < 10). For the mediation analysis, we report unstandardized regression coefficients. Path a (effect of age on strategy group) was the same for all four mediation analysis and resulted in a significant effect: $B = .10$, $p < .001$).

For pre-SMA/ACC, the effect of age on ROI activity (path c: $B = 0.13$, $p < .001$) was partly mediated by strategy group (path b: $B = 0.19$, $p = .007$; path c' : $B = .11$, $p < .001$; mediation effect (ab) = .02, $p = .01$ (95 % confidence interval (CI) = .004 - .036); proportion mediated (ab/c) = .15). In left DLPFC, the effect of age on ROI activity (path c: $B = 0.13$, $p < .001$) was also partly mediated by strategy group (path b: $B = 0.12$, $p = .031$; path c' : $B = .11$, $p < .001$; mediation effect (ab) = .01, $p = .04$ (95 % confidence interval (CI) = .0005 - .023); proportion mediated (ab/c) = .09). For right SPC, the effect of age on ROI activity (path c: $B = 0.21$, $p < .001$) was partly mediated by strategy group (path b: $B = 0.18$, $p = .033$; path c' : $B = .19$, $p < .001$; mediation effect (ab) = .02, $p = .04$ (95 % confidence interval (CI) = .001 - .037); proportion mediated (ab/c) = .09). A similar effect was found for left SPC, where the effect of age on ROI activity (path c: $B = 0.17$, $p < .001$) was partly mediated by strategy group (path b: $B = 0.18$, $p = .012$; path c' : $B = .15$, $p < .001$; mediation effect = 0.018, $p = .02$ (95 % CI: .004 - .034); proportion mediated = .11). In summary, we found that a significant portion of the variance in ROI activity in pre-SMA/ACC, left DLPFC and bilateral SPC was explained by strategy group, because strategy group partly mediated the relation between age and brain activity.

Whole brain analysis

In addition, we also tested if there were differences on a whole-brain basis between strategy groups, to investigate whether the ROI results were also observable on a whole-brain level, and to test whether there were effects of strategy group (above age) in areas outside of the ROIs. We used an ANOVA model with the four strategy groups as between-subjects variable and age as an additional regressor. We added age as a regressor because the strategy groups were not equally divided over age groups. First, we calculated the contrast Learning > Application for the four strategy groups separately, corrected for age (see Figure 6, Table 6). These analyses showed widespread activity (FWE-corrected at $p < .05$, > 10 contiguous voxels) for all strategy groups in the bilateral frontoparietal network, as well as in the pre-SMA/ACC, basal ganglia and occipital/temporal cortex. An F-test (with age-correction) was performed to see whether there was a main effect of strategy group. There were no significant clusters which survived FWE-correction.

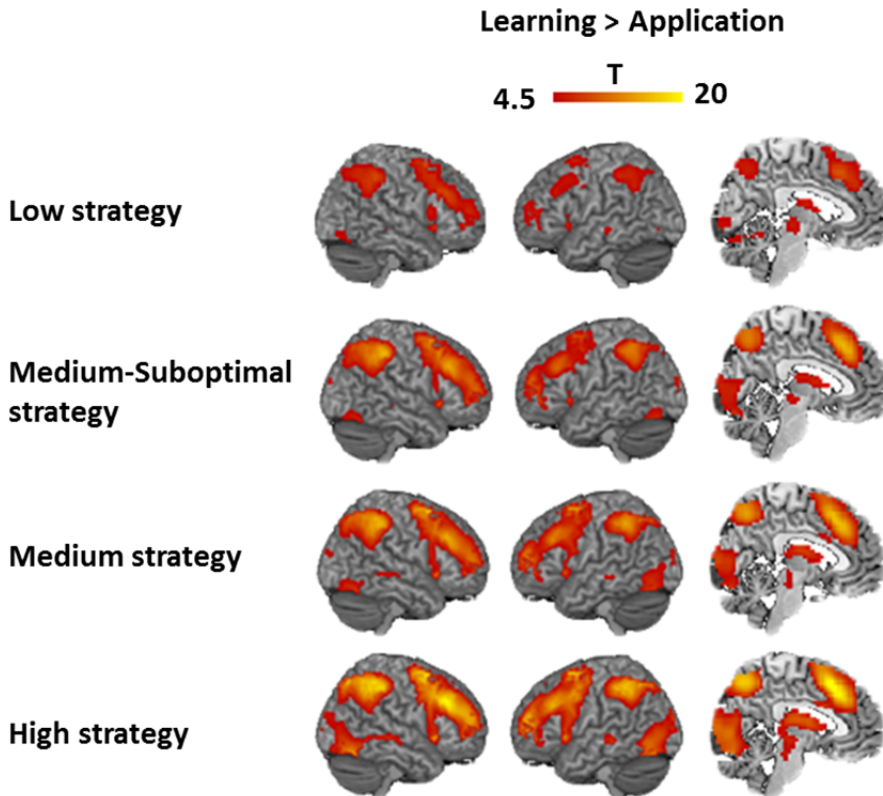


Figure 6: Areas showing activation for the contrast Learning > Application, separately for each strategy group (corrected for age), FWE corrected at $p < .05$, > contiguous 10 voxels.

Table 6: MNI coordinates local maxima activated for the contrast *Learning > Application* per strategy group, FWE corrected at $p < .05$, > contiguous 10 voxels

Strategy Group	Area of activation	x	y	z	voxels	T
Low	R inferior parietal lobule	45	-42	48	1778	10.39
	L inferior parietal lobule	-35	-45	42		7.54
	L inferior parietal lobule	-48	-39	45		7.39
	L superior medial gyrus	-3	24	45	2588	9.99
	R superior frontal gyrus	24	9	57		9.50
	R middle frontal gyrus	42	30	33		9.14
	R putamen	30	21	0	1018	3.06
	L insula lobe	-30	21	-3		8.43
	L caudate nucleus	-15	6	12		7.13
	L middle frontal gyrus	-30	51	18	277	6.93
	L middle frontal gyrus	-36	48	6		6.62
	L superior orbital gyrus	-30	54	-3		6.59
	R middle temporal gyrus	57	-30	-9	53	6.42
	L calcarine gyrus	-3	-90	-9	71	5.59
	R inferior occipital gyrus	42	-78	-12	36	5.16
	R fusiform gyrus	36	-75	-18		5.05
	R cerebellar vermis	6	-75	-18	21	5.11
	L cerebellum	-6	-78	-18		4.91
	R cerebellar vermis	3	-54	-15	27	5.04
Medium-SO	R inferior parietal lobule	45	-42	48	3087	17.34
	R angular gyrus	36	-57	45		16.69
	L inferior parietal lobule	-48	-42	48		12.67
	Superior medial gyrus	0	24	45	6258	16.17
	R putamen	30	21	0		14.77
	R middle frontal gyrus	45	27	36		14.67
	L superior occipital gyrus	-9	-99	6	972	9.09
	R cuneus	12	-96	12		8.21
	L fusiform gyrus	-39	-69	-18		7.78
Medium	R middle temporal gyrus	57	-30	-9	18	6.45
	R superior frontal gyrus	24	9	60	7784	18.16
	Superior medial gyrus	0	24	42		17.18
	R putamen	30	21	0		15.17
	R inferior parietal lobule	45	-42	48	5104	17.92

High	L inferior parietal lobule	-48	-45	48		15.11
	R precuneus	6	-66	51		14.13
	L middle temporal gyrus	-57	-30	-9	16	6.28
	R supplementary motor area	3	21	45	9890	22.07
	L supplementary motor area	-3	15	48		21.40
	L supplementary motor area	-6	9	54		20.57
	R supramarginal gyrus	39	-42	42	7336	21.61
	L inferior parietal lobule	-36	-48	45		18.52
	L inferior parietal lobule	-45	-42	45		18.52
	L middle temporal gyrus	-57	-33	-9	78	9.55

Note: Abbreviations: L = Left; R = Right

Age effects

Additionally, we investigated effects of age within our ANOVA model, to see which brain activation is related to age while controlling for strategy group. As can be seen in Figure 7, a positive relation with age (FWE-corrected, $p < .05$, > 10 contiguous voxels) was found in the frontoparietal network, pre-SMA/ACC, basal ganglia and occipital cortex (see also Supplementary Table 1), when controlling for strategy.

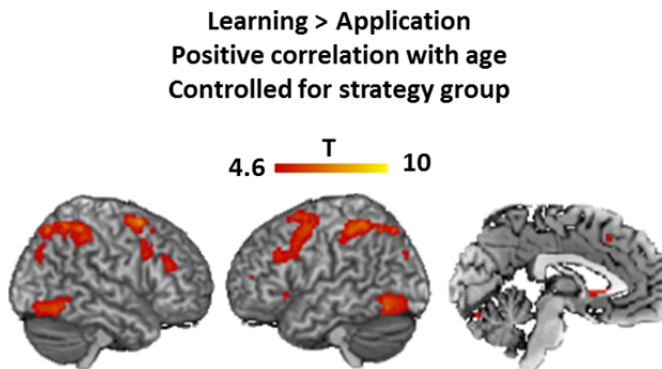


Figure 7: Areas showing activity for the contrast *Learning > Application* that correlated positively with age, while controlling for strategy group. Displayed at an FWE-corrected threshold at $p < .05$, > 10 contiguous voxels.

Discussion

An important question in research on executive functioning is whether developmental differences in behavior and neural activation can be explained by strategy differences. This study tested this hypothesis in a large sample of participants between ages 8-25 years in which statistical modeling approaches were combined with neuroimaging. The results showed that 1) learning from feedback in a sorting task resulted in variance in performance, which could be distinguished into four latent strategy groups, 2) even though age was strongly linked to strategy use, such that older participants were more often present in high performing strategy groups and young participants in low strategy groups, there was still considerable variance in strategy groups within age groups, 3) strategy explained additional variance in neural activity above age in pre-SMA/ACC, left DLPFC and bilateral SPC, and 4) there was still unique variance related to age differences in neural activity during feedback learning.

Underlying strategy groups in the feedback learning task

Consistent with prior developmental studies (Andersen et al., 2014; Raijmakers et al., 2001; Schmittmann et al., 2012, 2006), we found latent strategy groups that differed in performance on the feedback learning task. In the current feedback learning paradigm, participants were instructed to sort stimuli in one of three locations by using positive and negative feedback. Several strategies could be used to ensure optimal learning, such as focusing not only on feedback for the current stimulus, but also on feedback for the other two stimuli. In our sample, four different strategy groups could be distinguished. The highest performing strategy group responded almost perfectly in terms of the efficiency of strategy use. This strategy required optimal use of short-term memory, working memory and more complex executive functions (Diamond, 2013). The lowest performing strategy group, however, made more mistakes and did not adequately take into account the information from the other stimuli, thereby missing opportunities for learning. The lowest strategy seemed to involve regular flaws in STM, WM and more complex executive functions. The other two groups represented intermediate variants.

The division over strategy groups was related to age: Results indicated that younger children were more likely to be in a lower performing group, and older children and adults were more likely to be in a higher performing group. Still, there was considerable variation in strategy group within age groups: Some young children (8/9 years) belonged to the same strategy group as some of the adults, but there were also adults who were outperformed by young children. This supports the notion that it is important to study performance and strategy differences as opposed to age differences alone in learning tasks. The presence of different underlying strategies within age groups could make the interpretation of prior developmental studies that compared age groups difficult.

Neural differences between strategy groups

After discovering that different strategy groups could be distinguished in our sample, we investigated if these strategy groups explained developmental differences in neural activity. We used an fMRI contrast that revealed areas that are sensitive to feedback with learning value, i.e., areas that respond more to feedback that provides new information compared to feedback that provides information that is already known. We tested which areas showed a positive relation with age, and created ROIs based on this contrast for pre-SMA/ACC, DLPFC and SPC (see also Peters et al., 2014). We found that strategy group explained additional variance above age in pre-SMA/ACC, left DLPFC and bilateral SPC. In these areas, the highest performing group consistently showed more activation for these regions compared to the other groups. Moreover, with mediation analyses, we found that the proportion of the age effect on brain activity mediated by strategy was between 0.09 and .15. These mediated proportions were relatively low, especially compared to the study by Andersen et al. (2014) which used a similar approach but found stronger effects of strategy despite a much smaller sample size. This could be due to the fact that the task used in Andersen et al. was relatively more difficult, e.g., a majority of even the adult participants did not use the most optimal strategy. The relative contribution of age might be larger compared to strategy in a task such as in the current study where there is a clear developmental ‘end point’, i.e. most adults demonstrate optimal performance.

The finding that especially the highest performing group differed from the other groups in frontoparietal areas and the pre-SMA/ACC relates to prior studies which indicated that these areas are important for executive functioning. The pre-SMA/ACC is thought to be important to detect conflict between competing representations (Carter & van Veen, 2007) and for top-down control of response selection and preparation (Schulz, Bedard, Czarnecki, & Fan, 2011). The DLPFC and SPC are important for executive functions such as working memory (Klingberg et al., 2002), and the DLPFC is also associated with inhibition (Nyffeler et al., 2007) and cognitive flexibility (Ravizza & Carter, 2008). These executive processes are important to achieve the highest level of performance in the current feedback learning paradigm. Possibly, these fMRI results indicate that high performing learners are better at distinguishing feedback that is important for learning compared to relatively uninformative feedback (i.e., feedback during the application phase, which does not provide new information).

In addition, a whole-brain approach showed that the feedback learning network (pre-SMA/ACC, DLPFC, SPC and basal ganglia) was activated in all strategy groups (after age correction) while participants received feedback during learning compared to applying known rules. These findings are consistent with prior studies that showed that this network is implicated in feedback learning across development (Crone et al., 2008; van den Bos et al., 2009; van Duijvenvoorde et al., 2008). However, when testing for differences between the groups while correcting for age, there were no neural differences which survived correction for multiple comparisons. Together these results showed that age effects in pre-SMA/ACC, SPC and left DLPFC

during a feedback learning task are partly explained by strategy differences, although the effects of strategy may be relatively small compared to the effects of age.

Age versus strategy effects

Because of the relatively weak mediation effects of strategy on brain activity, we also investigated which brain areas show a positive correlation with age, while correcting for strategy group, to find unique activation related to age. We found widespread activation in the frontoparietal network, pre-SMA/ACC and basal ganglia, which survived correction for multiple-comparisons. Together with the ROI findings, these results suggest that age is the most significant contributor to neural activation patterns compared to strategy use in this study. This is consistent with the previously mentioned maturational viewpoint which suggests that age is a vital contributor to the development of patterns of neural activation (Dosenbach et al., 2010).

Therefore, our results only partly correspond to prior studies which found that age-related differences in neural activation can be explained by performance differences (Booth et al., 2004; Bunge et al., 2002; Koolschijn et al., 2011). For instance, in a longitudinal study, when performance and age were used as predictors for neural activation change in regions such as the DLPFC, SPC and pre-SMA/ACC, performance was a better predictor for changes in neural activity than age (Koolschijn et al., 2011). This highlights the need for longitudinal studies to investigate the development of learning and the underlying neural processes. Future research should investigate if participants progress to faster strategy groups with development, and if this is accompanied by more robust changes in neural activity than in this cross-sectional design. We propose that latent variable models provide a valuable method to detect performance-related versus age-related influences on neural activity.

Limitations

There are several limitations to this study. First, as is often the case in developmental studies, children received relatively more negative feedback compared to adults (e.g., Koolschijn et al., 2011). It is possible that our results were influenced by these differences in the amount of trials per feedback type, although these differences between age groups were relatively small. Second, even though the wide age range in this sample has many benefits, it was not possible to investigate effects of strategy in smaller age ranges, due to the unbalanced division of participants across strategy groups. A study with a similarly large sample but with less variation in age would be better suited to investigate effects of strategy within smaller age ranges. In addition, the study was cross-sectional and future studies should test changes in strategy use and neural activation patterns within the same individuals, to examine whether switching to a different strategy is accompanied by neural change within individuals. Finally, since we did not fit confirmative process models to the behavioral data, we cannot exclude the possibility that performance was not only organized in qualitatively different performance groups, but that there was also some con-

tinuous performance variation. However, there are good reasons to believe that qualitatively different strategy groups played an important role in performance differences of individuals. First, model comparisons indicated that the fit of a model that presumes continuous variation was relatively worse. Second, in the selected model the posterior probabilities to belong to a specific strategy group was for most participants either high (around 1.0) or low (around 0.0). Finally, strategy differences (as opposed to continuous performance variation) were found by studies that did fit confirmative models for feedback learning processes in other learning tasks (Raijmakers et al., 2001; Schmittmann et al., 2012, 2006).

Conclusion

In this study, we showed that in a feedback learning task, different underlying strategies could be detected within age groups, which were distinguishable at the neural level. These findings have important implications for traditional ways of analyzing developmental data. In future studies, it will be important to take into account individual differences in performance and strategy use, rather than comparing age groups alone. This research is informative in the context of unraveling the mechanisms underlying learning and learning difficulties and may contribute to interventions teaching children to adapt more efficient strategies to enable faster learning.

