



Universiteit
Leiden
The Netherlands

The neurocognitive development of social decision-making

Bos, W. van den

Citation

Bos, W. van den. (2011, April 12). *The neurocognitive development of social decision-making*. Retrieved from <https://hdl.handle.net/1887/16711>

Version: Not Applicable (or Unknown)

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

Downloaded from: <https://hdl.handle.net/1887/16711>

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

8. Striatum-medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning

Abstract

During development, children improve in learning from feedback to adapt their behavior. However, it is still unclear which neural mechanisms might underlie these developmental changes. In the current study we used a reinforcement learning model to investigate neurodevelopmental changes in the representation and processing of learning signals. Healthy volunteers between ages 8 and 22 (children: 8–11 years, adolescents: 13–16 years, and adults: 18–22 years) performed a probabilistic learning task while in a MRI scanner. The behavioral data demonstrated age differences in learning parameters with a stronger impact of negative feedback on expected value in children. Model-based analysis of imaging data revealed that the neural representation of prediction errors was similar across age groups, but prediction error-related functional connectivity between the ventral striatum and the medial prefrontal cortex shifted as a function of age, from stronger after negative feedback to stronger after positive feedback. Furthermore, the connectivity strength predicted the tendency to alter expectations after receiving negative feedback. These findings indicate that the underlying mechanisms of developmental changes in learning may not be related to differences in the computation of learning signals per se, but rather to differences in how learning signals are used to guide behavior and expectations.

8.1 Introduction

The ability to learn contingencies between actions and positive or negative outcomes in a dynamic environment forms the foundation of adaptive behavior (Rushworth & Behrens, 2008). Learning from feedback in probabilistic environments is sensitive to developmental changes, showing developmental improvements in learning from positive and negative feedback are observed until early adulthood (Crone & van der Molen, 2004; Hooper et al., 2004; Huizinga et al., 2006; van den Bos et al., 2009). Intriguingly, prior neuroimaging studies have demonstrated developmental differences in neural circuits associated with learning from feedback in a fixed, or static learning

environment (van Duijvenvoorde et al., 2008, Crone et al., 2008). These studies show that dorsolateral prefrontal cortex and parietal cortex are increasingly engaged when receiving negative feedback. However, in a probabilistic learning environment, learning is adaptive over trials and both positive and negative feedback informs future behavior. Therefore, an important question concerns the neural mechanisms that underlie developmental differences in adaptive probability learning.

A crucial aspect of adaptive learning is using feedback to estimate the expected value of the available options. The first step in estimating the expected value is the computation of prediction errors, that is, calculating the difference between expected and experienced outcomes. Prediction errors can be positive, indicating that outcomes are better than expected, or negative, indicating that outcomes are worse than expected (Sutton & Barto, 1998). Next, these prediction errors are used to update the expected value associated with the chosen option: the expected value increases when the prediction error is positive and decreases when the prediction error is negative.

Prior neuroimaging studies have shown that activity in the ventral striatum, a target area of dopaminergic midbrain neurons, correlates with positive and negative prediction errors (Knutson et al., 2000; Pagnoni et al., 2002; e.g. McClure et al., 2003; O'Doherty et al., 2003; McClure et al., 2004). The relation between prediction errors and subsequent learning is confirmed by studies demonstrating an association between the representation of prediction errors in the striatum and individual differences in performance on probabilistic learning tasks (Pessiglione et al., 2006; Schönberg et al., 2007). Recently, a developmental study revealed heightened sensitivity in the striatum to positive prediction errors in adolescents relative to children and adults (Cohen et al., 2010). Children (ages 8-12) did not show evidence for a prediction error signal in the striatum, whereas adolescents (ages 14-19) and adults (25-30) did. Therefore, it is possible that the representation of prediction errors is one mechanism contributing to the observed developmental changes in adaptive behavior.

Several neuroimaging studies have shown that activity in the medial prefrontal cortex (mPFC) correlates with the expected value of stimuli or actions (for review see Rangel et al., 2008). Representations of expected values in the mPFC are thought to be updated by means of fronto-striatal connections, relating striatal prediction errors to medial prefrontal representations (Houk & Wise, 1995; Pasupathy & Miller, 2005; Frank & Claus, 2006; Camara et al., 2009). In support of this hypothesis, recent studies have shown increased functional connectivity between the ventral striatum and mPFC during feedback processing (Camara et al., 2008; Munte et al., 2008). Furthermore, group

differences in learning may be related to the connectivity strength between the striatum and the PFC during feedback processing. For example, substance-dependent individuals have an intact striatal representation of prediction errors, but are impaired in subsequently using these signals for learning (Park et al., 2010). This study showed that there is a positive relation between learning speed and the strength of functional connectivity between the striatum and PFC (see also Klein et al., 2007). Therefore, a second possible mechanism that may contribute to developmental changes in adaptive behavior is an increase in striatal-mPFC connectivity. Indeed, there are also still substantial changes in anatomical connectivity between subcortical structures and the prefrontal cortex during adolescence (Supekar et al., 2009; Schmithorst & Yuan, 2010).

To test these two hypotheses, a computational model of reinforcement learning model was applied to investigate developmental differences in (a) the neural representation of prediction errors, and (b) changes in fronto-striatal connectivity. Participants of three age groups (children ages 8-11, adolescents ages 13-16 and young adults ages 18-22) performed a probabilistic learning task (Frank et al., 2004) in an MRI scanner. We expect that with age, there is an improvement in learning from probabilistic feedback (Crone & van der Molen, 2004; van den Bos et al., 2009). In order to capture age related changes in learning from positive and negative feedback separately, we use a reinforcement learning model with separate learning rates for positive and negative feedback (Kahnt et al., 2009). The individually estimated trial-by-trial prediction errors generated by this reinforcement model were subsequently used to test whether developmental differences in learning reflect functional differences in the representation of prediction errors or developmental changes in the propagation of prediction errors as measured by functional fronto-striatal connectivity (Park et al., 2010).

8.2 Material and Methods

8.2.1 Participants.

Sixty-seven healthy right-handed paid volunteers ages 8-22 participated in the fMRI experiment. Age groups were based on adolescent development stage, resulting in three age groups: children (8- to 11-year-olds, $n=18$; 9 female), mid-adolescents (13- to 16-year-olds, $n=27$; 13 female) and young adults (18- to 22-year-olds, $n=22$; 13 female). A chi square analysis indicated that gender distribution did not differ between age groups, $X^2(2) = .79$, $p = .67$. All participants reported normal or corrected-to-normal vision and participants or their caregivers indicated an absence of neurological or psychiatric impairments. Participants gave informed consent for the study and all

procedures were approved by the medical ethical committee of the Leiden University Medical Center.

Participants completed two subscales (similarities and block design) of either the Wechsler Adult Intelligence Scale (WAIS) or the Wechsler Intelligence Scale for Children (WISC) in order to obtain an estimate of their intelligence quotient (Wechsler, 1991, 1997). There were no significant differences in estimated IQ scores between the different age groups, $F(2, 66) = 1.63, p = .20$ (see Table 7.1, p. 132).

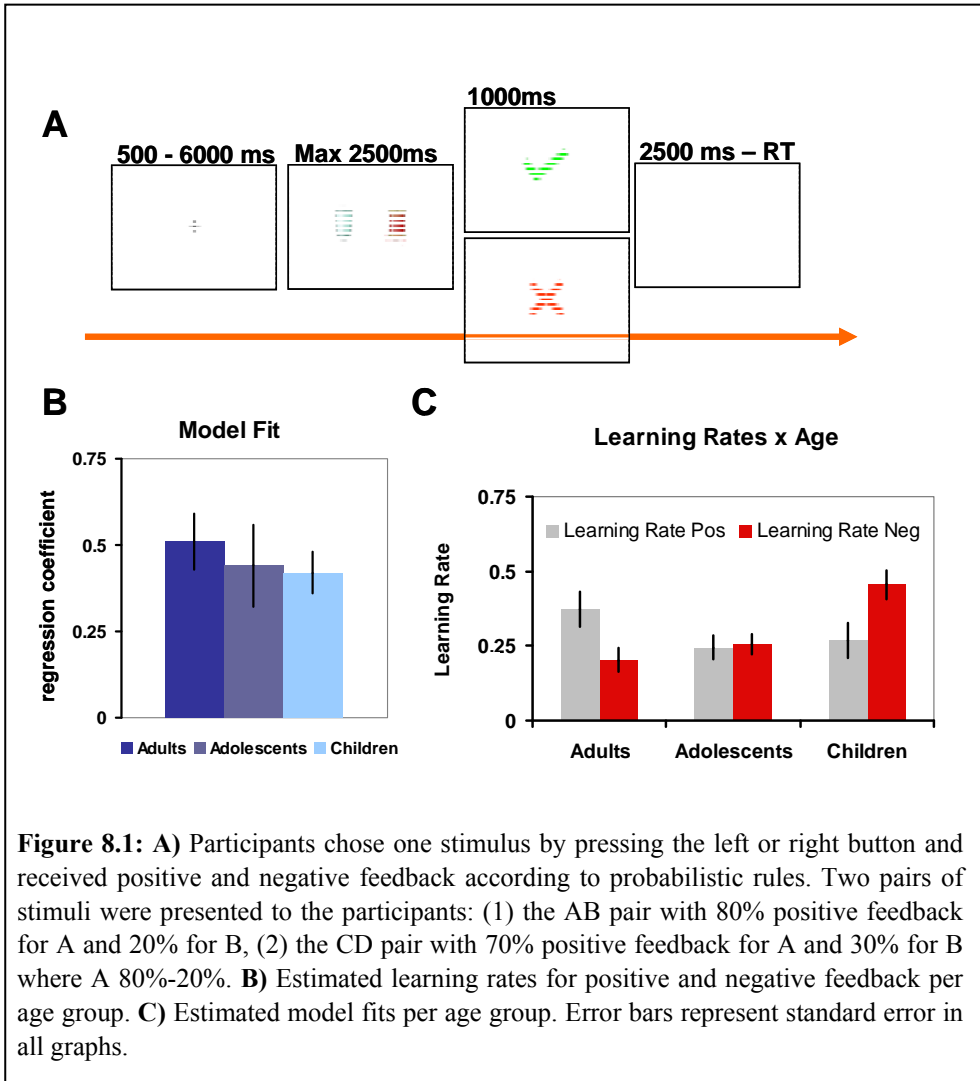
8.2.2 Task Procedure

The procedure for the probabilistic learning task (PLT, Frank et al., 2004; van den Bos et al., 2009) was as follows: The task consisted of two stimulus pairs (called AB and CD). The stimulus pairs consisted of pictures of everyday objects (e.g., a chair and a clock). Each trial started with the presentation of one of the two stimulus pairs and subsequently the participant had to choose one (e.g., A or B). Stimuli were presented randomly on the left or the right side of the screen. Participants were instructed to choose either the left or the right stimulus by pressing a button with the index or middle finger of the right hand. Responses had to be given within a 2500 ms window, which was followed by a 1000ms feedback display (see Figure 8.1 A). If no response was given within 2500 ms, the text “too slow” was presented on the screen.

Feedback was probabilistic; choosing stimulus A led to positive feedback on 80% of AB trials, whereas choosing stimulus B led to positive feedback on 20% of these trials. The CD pair procedure was similar, but probability for reward was different; choosing stimulus C led to positive feedback on 70% of CD trials, whereas choosing stimulus D led to positive feedback on 30% in these trials.

Participants were instructed to earn as many points as possible (as indicated by receiving a positive feedback signal), but were also informed that it was not possible to receive positive feedback on every trial. After the instructions and before the scanning session, the participants played 40 practice rounds on a computer in a quiet laboratory to ensure they understood the task.

In total, the task in the scanner consisted of two blocks of 100 trials each: 50 AB trials and 50 CD trials per block. The first and the second block consisted of different sets of pictures and therefore, participants had to learn a new mapping in both task blocks. The data from the last 60 trials of each block were also reported in another study using a rule-based analysis (van den Bos et al., 2009). The duration of each block was approximately 8.5 minutes. The stimuli were presented in pseudo-random order with a jittered interstimulus interval (min=1000 ms, max=6000 ms) optimized with OptSeq2 (Dale, 1999).



8.2.3 Reinforcement Learning Model

A standard reinforcement learning model (Sutton & Barto, 1998) was used to analyze behavioral and neural data (McClure et al., 2003; Cohen & Ranganath, 2005; Haruno & Kawato, 2006; Frank & Kong, 2008; Kahnt et al., 2008). The standard reinforcement learning model uses the prediction error (δ) to update the decisions weights (w) associated with each stimulus (in this case A, B, C or D) (Schultz et al., 1997; Holroyd & Coles, 2002). Thus, whenever feedback is better than expected, the model will generate a positive prediction error which is used to *increase* the decision weight of the chosen stimulus (e.g. stimulus A).

However, when feedback is worse than expected, the model will generate a negative prediction error, which is used to *decrease* the decision weight of the chosen stimulus (e.g. stimulus B). The impact of the prediction error is usually scaled by the learning rate (α). We extended the standard reinforcement learning model by using separate learning rates for positive feedback (α_{pos}) and negative feedback (α_{neg}) (e.g. Kahnt et al., 2008). Thus, positive and negative feedback might have a different impact of the decisions weights. To model trial-by-trial choices, we used the soft-max mechanism to compute the probability (p) of choosing a high probability target (A or C) on trial t as the logit transform of the difference in the decision weights in each trial (w_t) associated with each stimulus, passed through a biasing sigmoid function (Montague et al., 2004; Kahnt et al., 2008). For example, when stimulus pair AB is presented the probability of choosing A is determined by:

$$(1) \quad p(A)_t = \frac{e^{w(A)_t}}{e^{w(A)_t} + e^{w(B)_t}}$$

After each decision the prediction error (δ) is calculated as the difference between the outcome received ($r = 1$ for positive feedback and 0 for negative feedback) and the decision weight (w_t) for the chosen stimulus:

$$(2) \quad \delta_t = r_t - w(\text{chosen_stimulus})_t$$

Subsequently, the decision weights are updated according to:

$$(3) \quad w_{t+1} = w_t + \pi \times \alpha(\text{outcome})_t \times \delta_t$$

Where π is 1 for the chosen and 0 for the unchosen stimulus, $\alpha(\text{outcome})$ is a set of learning rates for positive (α_{pos}) and negative feedback (α_{neg}), which scale the effect of the prediction error on the future decision weights, and thus subsequent decisions. For example, a high learning rate for positive feedback but a low learning rate for negative feedback indicates that positive feedback has a high impact on future behavior, whereas negative feedback will hardly change future behavior. These two learning rates were individually estimated by fitting the model predictions ($p(\text{high probability stimulus})$) to participants' actual decisions. We used the multivariate constrained minimization function (fmincon) of the optimization toolbox implemented in MATLAB 6.5 for this

fitting procedure. Initial values for learning rates were $\alpha_{\text{pos}} = \alpha_{\text{neg}} = 0.5$ and for action values, $w(\text{left}) = w(\text{right}) = 0$.

8.2.4 Behavioral Analyses

To examine the correspondence between model predictions and participants' behavior, model predictions were compared with the actual behavior on a trial-by-trial basis. Model predictions based on estimated learning rates were regressed against the vector of participants' actual choices and individual regression coefficients were used to compare group differences in model fits. Differences in model fit between groups would indicate that other processes, for example a larger tendency to switch regardless of feedback, may play a relatively larger role in choice behavior in one group compared to the other. Only when there are no differences in model fit between groups one can confidently compare model parameters.

8.2.5 Data Acquisition

Participants were familiarized with the scanner environment on the day of the fMRI session through the use of a mock scanner, which simulated the sounds and environment of a real MRI scanner. Data were acquired using a 3.0T Philips Achieva scanner at the Leiden University Medical Center. Stimuli were projected onto a screen located at the head of the scanner bore and viewed by participants by means of a mirror mounted to the head coil assembly. First, a localizer scan was obtained for each participant. Subsequently, T2*-weighted Echo-Planar Images (EPI) (TR= 2.2 sec, TE= 30ms, 80 x 80 matrix, FOV = 220, 35 2.75mm transverse slices with 0.28mm gap) were obtained during 2 functional runs of 232 volumes each. A high-resolution T1-weighted anatomical scan and a high-resolution T2-weighted matched-bandwidth anatomical scan, with the same slice prescription as the EPIs, were obtained from each participant after the functional runs. Stimulus presentation and the timing of all stimuli and response events were acquired using E-Prime software. Head motion was restricted by using a pillow and foam inserts that surrounded the head.

8.2.6 fMRI Data Analysis

Data were preprocessed using SPM5 (Wellcome Department of Cognitive Neurology, London). The functional time series were realigned to compensate for small head movements. Translational movement parameters never exceeded 1 voxel (< 3 mm) in any direction for any subject or scan. There were no significant differences in movement parameters between age groups $F(2, 65) = .15, p = .85$. Functional volumes were spatially normalized to EPI templates.

The normalization algorithm used a 12 parameter affine transformation together with a nonlinear transformation involving cosine basis functions and resampled the volumes to 3 mm cubic voxels. Functional volumes were spatially smoothed using a 8 mm full-width half-maximum Gaussian kernel. The MNI305 template was used for visualization and all results are reported in the MNI305 stereotaxic space (Cosoco, Kollokian, Kwan, & Evans, 1997)

Statistical analyses were performed on individual participants' data using the general linear model in SPM5. The fMRI time series data were modeled by a series of events convolved with a canonical haemodynamic response function (HRF). The presentation of the feedback screen was modeled as 0-duration events. The stimuli and responses were not modeled separately as these occurred in one prior or overlapping EPI images as feedback presentation. To investigate the neural responses to feedback valence, independent of learning conditions, we set up a general linear model (GLM) with the onsets of each feedback type (positive and negative) as regressors.

To examine the neural correlates of reward prediction errors, we set up a second GLM with a parametric design. In this model, the stimulus functions for feedback were parametrically modulated by the trial-wise prediction errors derived from the reinforcement learning model. The modulated stick functions were again convolved with the canonical HRF. These regressors were then orthogonalized with respect to the onset regressors of positive and negative feedback trials and regressed against the BOLD signal.

Finally, to investigate age linear and quadratic age trends we applied polynomial expansion analysis (Büchel et al., 1996) with age as continuous variable, using the forward model selection as described by Büchel and colleagues (1998). Thresholds were set to $p < .001$ uncorrected for the whole group analyses, with an extend threshold of 15 continuous voxels (cf. Kahnt et al., 2008). We used the Marsbar toolbox for use with SPM5 (<http://marsbar.sourceforge.net>, Brett et al. 2002) to perform Region of Interest (ROI) analyses to further characterize patterns of activation and estimate individual differences in connectivity measures.

8.2.7 Functional Connectivity Analyses

To explore the interplay between the ventral striatum and other brain regions during reinforcement-guided decision-making, functional connectivity was assessed using psychophysiological interaction (PPI) analysis (Friston, 1994; Cohen et al., 2005; Cohen et al., 2008). The functional whole brain mask, in which activity correlated significantly with prediction errors for the whole group, was masked with an anatomical striatum ROI of the Marsbar toolbox that included the bilateral caudate, putamen and nucleus accumbens, to create

the seed region of interest (ROI). The method used here relies on correlations in the observed BOLD time-series data and makes no assumptions about the nature of the neural event that contributed to the BOLD signal (Cohen et al., 2008). For each model, the entire time series over the experiment was extracted from each subject in the clusters of the (left and right) ventral striatum. Regressors were then created by multiplying the normalized time series of each ROI with condition vectors that contained ones for four TRs after positive or negative prediction errors and zeros otherwise (see also Cohen & Ranganath, 2005; Kahnt et al., 2008; Park et al., 2010). Thus, the two condition vectors of positive and negative prediction errors (containing ones and zeros) were each multiplied with the time course of each ROI. These regressors were then used as covariates in subsequent analyses.

The time series between the left and right hemispheres for the ventral striatum were highly correlated (averages across runs and participants were $r = .84$). Therefore, parameter estimates of left- and right structures were collapsed, and thus, represent the extent to which feedback-related activity in each voxel correlates with feedback-related activity in the bilateral ventral striatum.

Individual contrast images for positive vs. negative prediction errors were computed and entered into second-level one-sample t-tests. In order to find age related differences in the whole-brain analyses of functional connectivity with the ventral striatum, we performed a regression analyses with an additional regressor for age. Thresholds for the connectivity analyses were also set to $p < .001$ uncorrected, with an extend threshold of 15 continuous voxels.

8.3 Results

8.3.1 Behavioral data

Reinforcement learning. First, we assessed how the model parameters differed between age groups. First of all, there was a good fit of the model to participants' behavior; the average regression coefficient was significantly above zero for all age groups (all p 's $< .001$, Figure 8.1 B). Importantly, the model fit did not differ significantly between groups ($F_{(2,64)} = .96$, $p = .38$), reassuring that parameters estimations could be compared between groups.

Next, a two (learning parameters) x three (age groups) ANOVA tested for age differences in learning from positive and negative feedback. This analysis showed a significant group by parameter interaction ($F_{(2,64)} = 12.34$, $p < .001$, see Figure 8.1 C), post-hoc tests revealed that there was an age related decrease in α_{neg} , $F_{(2,67)} = 9.87$, $p < .001$, and a marginal age related increase in α_{pos} , $F_{(2,67)} = 2.96$, $p = .06$. This result indicates that the relative influence of positive

feedback on expected values decreased with age and the relative influence of negative feedback on expected values increased with age.

8.3.2 fMRI results

Model-based fMRI. Across all participants, individually generated trial-wise prediction errors (positive and negative combined) correlated with BOLD signal in bilateral ventral striatum, MPFC, posterior anterior cingulate cortex (pACC) and the bilateral amygdala extending into the parahippocampal gyrus (Figure 8.2 A, and Table 8.1). Activity in the ventral striatum was localized at an area comprising the ventral intersection between the putamen and the head of the caudate. Tests for positive and negative prediction errors separately revealed comparable results.

Whole brain regression analyses for age differences revealed no linear or non-linear age group differences (Figure 8.2 B). This analysis was repeated for positive and negative prediction errors separately and these analyses also revealed no linear or non-linear age effects. This finding shows that the prediction error (positive or negative) is not represented differently between the three age groups.

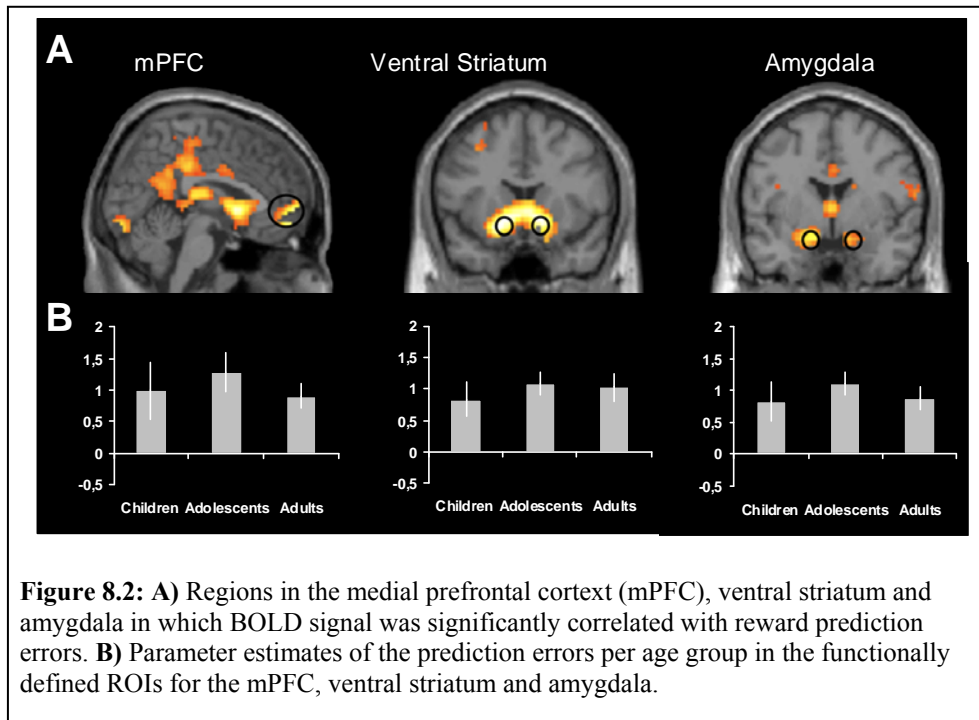


Table 8.1 : Brain Regions revealed by whole brain contrasts.

Anatomical region	L/R	BA	Z	MNI coordinates		
				x	y	z
Positive > Negative Feedback						
Ventral Striatum	L/R		7.49	-6	12	-3
Dorsolateral prefrontal cortex	L	8	4.61	-27	24	51
Superior parietal cortex	L	7	4.23	-30	-75	48
Precuneus	L/R	31	4.07	-3	-36	33
Medial PFC	L/R	10/11	4.03	3	54	-12
Visual Cortex	L/R	17	4.50	27	-93	-9
Negative > Positive Feedback						
Dorsal Anterior Cingulate Cortex	L/R	32	4.43	9	21	36
Prediction Error						
Ventral Striatum (caudate & putamen)	L/R		6.29	-6	9	3
Left Amygdala	L/R		5.50	-12	3	-18
Right Amygdala	R		5.05	18	6	-18
Medial PFC	L/R	10/11	5.84	0	54	3
Posterior Cingulate Cortex	L/R	32	4.83	0	-33	41
Visual Cortex	L/R	17	6.63	-18	-93	-18
PPI (positive > negative)						
Medial Prefrontal Cortex	L/R	10	5.47	-4	40	6
Pre-SMA	R	6	4.98	9	30	57
Right Anterior Insula / IFG	R		4.46	41	23	-9
Left Anterior Insula / IFG	L		4.67	-44	21	-3
Ventral Striatum (caudate & putamen)	L/R		7.50	9	9	3
PPI (positive > negative) x Age						
Medial PFC	L	10	4.02	-8	45	10

MNI coordinators for main effects, peak voxels reported at $p < .001$, at least 20 contiguous voxels.

Functional Connectivity. Functional connectivity between the striatum and other brain regions was assessed during processing of negative and positive feedback using PPI. The contrast used for testing functional connectivity was positive > negative feedback. Note that the vectors for positive feedback events contain all positive prediction error events, and the vectors for negative

feedback events contain all negative prediction error events. Enhanced functional connectivity was found during positive > negative feedback between the bilateral ventral striatum seed and the mPFC (Figure 8.3 A), dACC, pre-SMA, and bilateral anterior Insula extending into the inferior frontal gyrus. The opposite contrast (negative > positive feedback) did not reveal any significant changes in functional connectivity.

Next, we examined age differences in ventral striatum connectivity by adding age as a regressor to the whole-brain PPI analysis. These analyses revealed age related increases in functional connectivity of the ventral striatum seed with the mPFC (BA10) for positive > negative feedback (Figure 8.3 B). No other areas were found when testing for non-linear age effects in functional connectivity.

To further illustrate the age related changes in fronto-striatal connectivity we extracted the strength of functional connectivity between ventral striatum and mPFC for each participant and plotted it against age as a continuous variable (Figure 8.3 C). This plot reveals that the connectivity pattern shifts from a stronger connection after negative feedback for the youngest participants towards a stronger connection after positive prediction errors for the oldest participants.

Finally, we performed ROI analyses to investigate whether striatum-mPFC connectivity was related to the individual learning parameters. The differential connectivity strength (positive > negative) between the ventral striatum and mPFC ROI was used to predict the individual differences in learning rates for positive and negative feedback. The relative connectivity measure correlated negatively with the learning rate for negative feedback ($r = -.39$, $p < .001$, Figure 8.3 D), and moderately positively with the learning rate for positive feedback ($r = .23$, $p = .07$). Thus, there was stronger striatum-mPFC coupling during negative > positive feedback in participants for whom negative feedback had a relatively large impact on future expected value, whereas the reverse was true (i.e., stronger coupling during positive > negative feedback) in participants for whom positive feedback had a relatively large impact on future expected value.

To summarize, increased functional connectivity between the ventral striatum and mPFC was observed during processing of positive feedback compared to negative feedback. Furthermore, this analysis revealed that the relative strength in striatum-mPFC connectivity correlated positively with age, but negatively with the learning rate for negative feedback.

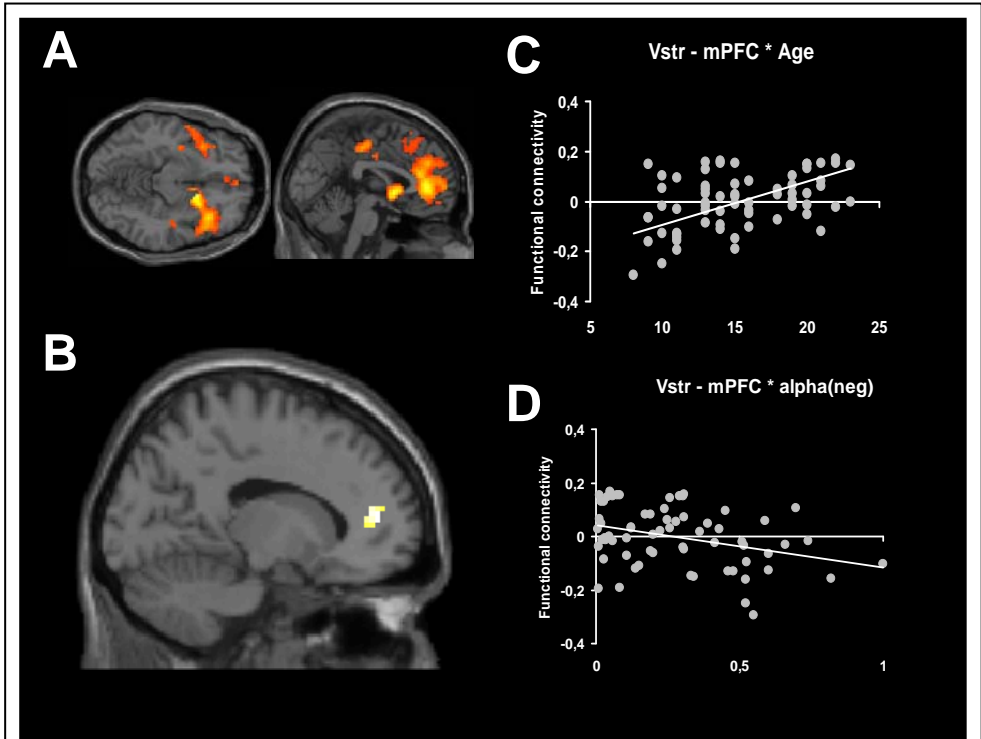


Figure 8.3: A) Regions which showed increased functional connectivity with the striatal seed region after positive compared to negative feedback. (B) Region in the mPFC that revealed age related changes in functional connectivity with the striatal seed region. Both statistical maps are all thresholded at $p < .001$, uncorrected, $k = 15$. (C) Scatterplot depicting the relationship between the functional connectivity measure of the striatum-mPFC (positive > negative feedback) and age. (D) Scatterplot depicting the relationship between the functional connectivity measure of the striatum-mPFC (positive > negative feedback) and learning rate (α_{neg}).

8.4 Discussion

The goal of this study was to examine developmental changes in the neural mechanisms of probabilistic learning. The reinforcement model showed that with increasing age, negative feedback had decreasing effects on future expected values. Imaging analyses revealed that ventral striatum activation following prediction errors did not differ between age groups; however, age differences in the learning parameters were associated with an age related increase in functional connectivity between ventral striatum and the mPFC. These behavioral data and their neural correlates allow a deeper understanding of how children, adolescents and adults learn in a changing environment. The discussion will be organized according to these themes.

Developmental changes in learning parameters

Using a reinforcement learning model we were able to disentangle differences in sensitivity to positive and negative feedback by estimating learning rates for positive and negative feedback separately. These estimated learning rates reflect the degree to which the future expected value of a stimulus will be changed after positive or negative prediction errors. As expected, the model-based analyses of learning behavior showed that with age there is a decrease in the learning rate for negative prediction errors (α_{neg}). This finding indicates that with increasing age, the impact of negative prediction errors on the future expected value decreases. These results are consistent with developmental studies that have shown that adults are less influenced by irrelevant negative feedback (Crone et al., 2004). Furthermore, compared to younger adults, older adults have been shown to report less negative arousal to anticipated losses (Samanez-Larkin et al., 2007). Taken together, these results show that the current reinforcement model can capture the subtle age related changes in adaptive learning, and thus provides a solid basis for exploring the underlying neurodevelopment changes in representing and the processing of learning signals.

Neural Representation of prediction errors

Consistent with previous studies, trial-by-trial prediction errors generated by the reinforcement learning model correlated with activity of a network of areas including the ventral striatum, mPFC and the amygdala (Pagnoni et al., 2002; McClure et al., 2003; O'Doherty et al., 2003; Cohen & Ranganath, 2005). This result indicates that these areas are sensitive to differences in expected vs. received feedback; showing increased activation when feedback is better than expected and decreased activation when the feedback is worse than expected. Our analyses did not reveal any (linear or non-linear) age related differences in prediction errors (positive or negative). These findings are consistent with prior studies using cognitive learning tasks, which have also reported early maturation of subcortical regions and protracted development of cortical brain areas (Casey et al., 2004; van Duijvenvoorde et al., 2008; Velanova et al., 2008). However, a recent developmental study of reward-based learning using a comparable reinforcement model, with a single learning rate (for both negative and positive feedback), has shown heightened sensitivity to positive prediction errors in adolescents compared to children and adults (Cohen et al., 2010). It should be noted however, that Cohen and colleagues compared different age groups, as adolescence in this study was defined as the age range 14-19 years, and adulthood as 25-30 years. In this respect, the findings of the current study and the findings of Cohen et al. are not directly comparable. In future studies, it

will be important to test for changes in prediction errors across a wider age range and differentiating between different phases of adolescence.

The results of the current study provide different findings in comparison to affective paradigms. These studies have reported increased sensitivity of the striatum in adolescence after receiving monetary rewards or highly emotional stimuli (Galvan et al., 2006; McClure-Tone et al., 2008; Van Leijenhorst et al., 2009), which may trigger the peak in adolescent reward processing. Interestingly, Cohen et al. (2010) observed adolescent-specific increases in reaction times for 25 cents relative to 5 cents rewards. In future studies, it will be important to examine whether the prediction error representation can be modulated by the use of affective tasks or reward manipulations, and whether these effects are dependent on the development of the dopaminergic system during adolescence (for a review see Galvan, 2010).

Developmental changes in striatum-mPFC connectivity

Connectivity analyses revealed that during feedback processing the seed region in the ventral striatum sensitive to prediction errors showed increased functional connectivity with the mPFC, pre-SMA, and bilateral anterior insula/IFG during positive compared to negative feedback. This pattern of connectivity is consistent with several studies that have shown feedback-related changes in functional connectivity of the striatum (for a review see Camara et al., 2009).

Subsequent analyses revealed age related changes in striatum–mPFC functional connectivity. The pattern shifted from stronger connectivity after negative feedback for the youngest participants towards stronger connectivity after positive feedback for the oldest participants. This suggests that shifts in feedback-dependent striatum-mPFC connectivity may underlie developmental changes in learning behavior. This interpretation is in line with an adult study which has shown that the strength of ventral striatum-mPFC connectivity following feedback is related to the adjustment of behavior on subsequent trials (Camara et al., 2008). This hypothesis is further supported by the correlation between striatum-mPFC connectivity and estimated learning rate parameter for negative prediction errors in the current study.

Given that during adolescent development there are still substantial changes in structural connectivity within the prefrontal cortex (Schmithorst & Yuan, 2010) it could be hypothesized that the developmental differences in striatum-mPFC functional connectivity are related to changes in structural connectivity between these two structures (Cohen et al., 2008). In future developmental studies, it will be of interest to combine measures of structural and functional connectivity in order to further explore this hypothesis.

Additionally, it should be noted that the functional connectivity measure is uninformative about the directionality of the influence between different brain regions (Friston, 1994). Applying methods such as structural equation modeling and dynamic causal modeling (Friston et al., 1997), which take directionality into account, could further increase our knowledge of the underlying mechanisms of developmental changes in adaptive learning.

A final question concerns how these results relate to previous developmental studies of feedback processing in static environments (van Duijvenvoorde et al., 2008; Crone et al., 2008). Learning theories have suggested two separate systems that operate in parallel during feedback learning (Dickinson & Balleine, 2002); one system that operates on task explicit representations, such as rules, and another system that operates on statistical contingencies of the environment, such as feedback probabilities. Recently, a study showed that updating task representations relies on the DLPFC-parietal network, whereas updating feedback probabilities was associated with the striatum (Gläscher et al., 2010). Thus, it is likely that developmental changes in the DLPFC-parietal network represent differences in the learning system that operates on rule-based task representations, whereas the current study shows developmental differences in the system tracking statistical contingencies (see also Galvan et al., 2006; Cohen et al., 2010). The challenge for future developmental studies will be to disentangle the relative contributions of these networks, and to understand how these two systems contribute to developmental changes in feedback learning.

Conclusion

Previous studies have shown that either changes in the representation of the prediction errors in the striatum (Schönberg et al., 2007) or the connectivity of the ventral striatum with the prefrontal cortex (Klein et al., 2007; Park et al., 2010) are related to individual differences in feedback learning. In the current study we provide evidence that developmental differences in feedback learning may not be due to differences in the representation of the prediction errors per se, but rather to developmental changes in the functional connectivity between the striatum and the mPFC. This finding suggests that children do not differ in their ability to track the statistical contingencies in the task, but rather process the learning signals differently. These findings advance our understanding of the neurodevelopmental underpinnings of probabilistic learning and highlight the importance of studying neural circuits in addition to specific brain regions (Camara et al., 2009).