Psychophysiology, 46 (2009), 1043–1053. Wiley Periodicals, Inc. Printed in the USA.

DOI: 10.1111/j.1469-8986.2009.00838.x

Developmental differences in learning and error processing: Evidence from ERPs

BEN EPPINGER, BARBARA MOCK, AND JUTTA KRAY
Department of Psychology, Saarland University, Saarbrücken, Germany

Abstract

This study examined developmental differences in the ERP correlates of internal and external error processing (ERN and FRN) during learning. A probabilistic learning task was applied in which feedback validity was manipulated. The behavioral data showed similar accuracy for children and adults when feedback was valid, whereas age differences were obtained when it was partially invalid. We found no reduction of the ERN for children compared to adults when performance levels were equated. Yet, contrary to adults, children did not differentiate between responses when feedback was partially invalid, indicating that they are less able to represent the correctness of a response when there is interference during learning. Moreover, we found a larger FRN and reduced ERP learning effects for positive feedback for children, suggesting that they are more sensitive to external error feedback and less able to disengage from positive feedback during learning.

Descriptors: Development, Reinforcement learning, Error processing, ERN

Reinforcement learning plays a central role during cognitive development, and it appears to be particularly important for the flexible acquisition of behavior in sensitive periods during childhood and adolescence (see Johnson & Munakata, 2005). Behavioral studies on developmental differences in reinforcement learning have shown that especially younger children are impaired in learning when they have to adapt to changes in reward frequency, or when they have to build up predictions on sequential effects during learning (Offenbach, 1964; Stevenson & Weir, 1959).

The physiological processes that underlie reinforcement learning have been extensively studied using electrophysiological approaches in nonhuman primates (see Schultz, 2002). One of the central findings of these studies was that the midbrain dopamine system plays a key role for reinforcement learning by providing learning signals that are used by other brain areas to flexibly adapt behavior (Montague, Hyman, & Cohen, 2004). However, little is known about the ontogenetic development of the dopamine system. Indirect evidence for the role of dopamine during development comes from event-related potential (ERP) studies on error processing in children with attention-deficit hyperactivity disorder (ADHD). ADHD is a developmental disorder that has been associated with alterations in the dopamine system and its projections to the anterior cingulate cortex (ACC) (see Biederman & Faraone, 2005). Results of these studies suggest that ADHD children are impaired in the processing of internal and external error information and that this impairment might result from a reduced dopaminergic input to the ACC (Liotti, Pliszka, Perez, Kothmann, & Woldorff, 2005; van Meel, Oosterlaan, Heslenfeld, & Seargent, 2005).

The reinforcement learning (R-L) theory by Holroyd and Coles (2002) integrates the role of dopamine for learning with the error processing system associated with the ACC. The theory assumes that the ERP correlates of internal and external error processing (ERN and FRN, respectively) are generated when a negative reinforcement learning signal from the dopamine system is conveyed to the ACC. Hence, the ERN/FRN is proposed to reflect a negative prediction error that is generated when the outcome of an action is worse than expected. The ERN (Gehring, Goss, Coles, Meyer, & Donchin, 1993) or Ne (Falkenstein, Hohnsbein, & Hoormann, 1995) is a negative ERP component that is elicited around 80 ms after an erroneous response. It is typically maximal at fronto-central electrodes and is assumed to be generated in the dorsal ACC (Holroyd et al., 2004; van Veen & Carter, 2002). Similar to the ERN, the feedback-related negativity (FRN) shows a medial frontal topography but is elicited by external error feedback (Gehring & Willoughby, 2002; Müller, Möller, Rodriguez-Fornells, & Münte, 2005; Nieuwenhuis, Yeung, Holroyd, Schurger, & Cohen, 2004). The FRN has been first observed by Miltner, Braun, and Coles (1997) and can be found between 200 and 300 ms after the onset of a negative feedback stimulus. Holroyd and Coles (2002) showed that consistent with the R-L theory the ERN increased with learning

This work was funded by the Deutsche Forschungsgesellschaft (grant SFB 378). We are grateful to Philippe Ginsbourger for help during data acquisition and Axel Mecklinger for insightful comments and discussion. Furthermore, we thank our participants for their contribution to the study.

Address reprint requests to: Ben Eppinger, Princeton Neuroscience Institute, Green Hall, Princeton University, Princeton, NJ 08540, USA. E-mail: eppinger@princeton.edu

in a probabilistic learning task. This suggests that the more the participants learn to predict the correctness of their response, the larger the ERN. In contrast, the FRN decreased with learning, indicating that participants rely less on external error feedback with learning.

Results of recent developmental studies on internal error processing showed that the ERN increases with age until late adolescence (Davies, Segalowitz, & Gavin, 2004; Ladouceur, Dahl, & Carter, 2007; Santesso, Segalowitz, & Schmidt, 2006). Furthermore, these changes seem to follow a nonlinear trajectory, showing the largest developmental differences between 10 and 13 years of age (Davies et al., 2004). However, in all of these studies age differences in the ERN were accompanied by age differences in performance measures, suggesting that developmental changes in the ERN might have been confounded by differences in task performance. To address this question, Hogan, Vargha-Khadem, Kirkham, and Baldeweg, (2005) compared adolescents (12–18 years) to adults (18–22 years) using forced-choice visual reaction-time tasks of different complexity. A reduced ERN and behavioral impairments for adolescents were only observed in the more complex task version, indicating that the control of task performance is critical when comparing the ERN between age groups (see also Eppinger, Kray, Mock, & Mecklinger, 2008).

The ERN is not the only ERP component that is specific to errors. Most of the aforementioned studies have also investigated a second ERP correlate of internal error processing, the so-called error positivity (Pe). The Pe follows the ERN between 200 and 500 ms post-response and is typically maximal at central electrodes (Overbeek, Nieuwenhuis, & Ridderinkhof, 2005). The Pe has been proposed to reflect either the conscious recognition of an error (Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001) or its motivational significance (Overbeek et al., 2005). In contrast to the ERN, which has been suggested to be reduced in children, the Pe has been shown to be unaffected by age.

The literature on age differences in the processing of external error feedback during learning is rather scarce. To our knowledge, two studies have so far investigated the FRN in children (Holroyd, Baker, Kerns, & Mueller, 2008; van Meel et al., 2005). However, these studies focused on the effects of ADHD on the FRN and did not include an adult control group, which makes it difficult to draw conclusions about developmental differences in the FRN. In a recent developmental study, Crone, Jennings, and van der Molen (2004) examined changes in heart rate to negative feedback during probabilistic learning in 8–12-year-old children and adults. They showed that adult heart rate was slowed following negative feedback in a learning condition, but not in a non-learning condition. In contrast, heart rate in children was slowed for both conditions, indicating that children are more sensitive to negative feedback and less able to distinguish informative from uninformative feedback during learning.

The present study aims to provide insight into the question whether children differ from adults in the ability to use negative reinforcement learning signals for learning and whether this is reflected in the ERP correlates of internal and external error processing (ERN and FRN, respectively). We applied a probabilistic learning task in which we manipulated feedback validity. In the 100% validity condition, feedback was always valid. In the 80% validity condition, feedback was valid in 80% of the trials and invalid in 20% of the trials. Thus, participants were less certain about the feedback and by this less able to learn. In the 50% condition, feedback was delivered randomly and no learning was possible. To avoid potentially confounding effects of performance differences between age groups, we used a procedure that adaptively adjusted the response deadline based on the number of time-outs (see Methods). The idea of this procedure was to give children more time for responding, which should reduce the number of time-outs (trials in which they received no feedback) and by this support their ability to learn.

We expected that the deadline procedure should enable children to learn comparably to adults, at least in the 100% validity condition. However, similar to older adults (see Eppinger et al., 2008), we expected that children would be impaired in accuracy when invalid feedback interferes with learning in the 80% condition. Based on recent findings (Davies et al., 2004; Ladouceur et al., 2007; Santesso et al., 2006), we expected impaired internal error processing in children as reflected in a reduced ERN during learning. However, given the findings by Eppinger et al. (2008) and Hogan, Vargha-Khadem, Kirkham, and Baldeweg (2005), it is an open question whether we still find age differences in the ERN when accuracy levels are equated between age groups. In contrast to the ERN, the existing data suggest that no age differences should be found in the error positivity (Pe) (Davies et al., 2004; Ladouceur et al., 2007; Santesso et al., 2006). According to the studies by Crone et al. (2004) and van Meel et al. (2005), we expected that children would be less able to disengage from external error feedback and show increased FRNs for all validity conditions and no change of the FRN with learning.

Methods

Participants

Twenty-one adults[1] and 21 children participated in the study and received 22.5 Euro for participation. One younger adult had to be excluded due to technical problems. Two younger adults were excluded because they did not commit enough error trials. One child felt so uncomfortable with the EEG setting that the experimental session was cancelled. Three children were excluded because they performed at chance level in the 100% condition ($M = 0.48$, $SD = 0.03$) and responded much faster ($M = 278$ ms, $SD = 22$) than the mean of the children group ($M = 404$ ms, $SD = 128$). The effective sample consisted of 18 adults between 19–24 years of age (mean age = 20.8 years, $SD = 1.8$, 9 female) and 17 children between 10–12 years (mean age = 11.4 years, $SD = 0.8$, 9 female). According to self-report, participants were healthy, had a right-hand preference, no color blindness, no history of neurological or psychiatric problems, and were known to have no learning disorder.

Stimuli and Task

The stimulus set consisted of 36 colored images of objects (Snodgrass & Vanderwart, 1980) and belonged to one of six categories (clothes, vehicles, fruit, vegetables, furniture, domestic appliances). Participants were asked to make a two-choice decision upon presentation of the stimulus and to press one of two response keys (C and M on a computer keyboard). They were told to infer the stimulus-response mappings by trial and error. The German words 'RICHTIG' ('correct') printed in green and 'FALSCH' ('incorrect') printed in red served as feedback stimuli.

[1]The sub-sample of younger adults already served as a control group in a study on the effects of aging on learning and error processing.

When the response deadline was missed, the words 'ZU LANGSAM' ('too slow') were presented. For the children, we constructed a cover story similar to a Donald Duck comic. They were told that they should use the feedback to help Scrooge McDuck to sort objects into two safes (the two response buttons) to protect them from the "Beagle Boys." To increase motivation, participants could win up to 450 Euro Cents per block, depending on their performance in the 100% condition. At the end of each block, they received feedback about the money they had won. To further motivate the children, there was a break of 15 seconds in the middle of each block in which a performance-independent monetary feedback was displayed (always less than at the end of the block). For adults the outcomes at the end of the block ranged from 50 to 450 Euro Cents. For children they ranged from 250 to 450 Euro Cents (the lower value was clipped because of the feedback in the middle of the block).

Experimental Design

The design involved three learning conditions in which we manipulated feedback validity. In the 100% validity condition, one stimulus (A) was mapped to the right response key and the other stimulus to the left response key (B). If participants pressed the right key to stimulus A, they always received positive feedback, whereas they received negative feedback when they pressed the left key (and vice versa for B). Two other stimuli (C and D) were associated with the 80% condition. If participants pressed the left key to stimulus C, they received positive feedback in 80% and negative feedback in 20% of the key presses. If they pressed the right key, they received negative feedback in 80% and positive feedback in 20% of the key presses (and vice versa for D). In the 50% condition (stimuli E and F), feedback was delivered randomly. Each block involved a new set of six stimuli, which were drawn randomly (without replacement) from the six categories.

Procedure

At the beginning of the experiment, each participant (the parents, in case of the children) filled out an informed consent and a demographic questionnaire. Then they performed the experimental task, which consisted of one practice block and five experimental blocks. In the practice block (150 trials), the participants were familiarized with the experimental setting. In the five experimental blocks, each of the six imperative stimuli were presented 50 times in random order. Each trial started with the presentation of a fixation cross (500 ms), which was followed by the imperative stimulus (500 ms). After the key press, a blank screen was displayed (500 ms) and then the feedback appeared (500 ms). For adults, the response deadlines ranged between 600 and 1000 ms. For children, we applied deadlines between 800 and 1200 ms to account for their slower response times. Participants started with a deadline of 800 ms. After the first trial, the procedure kept track of the proportion of time-outs (time-outs relative to trials performed). If the proportion of time-outs was smaller than 2%, a deadline of 600 ms (adults) or 800 ms (children) was applied. With steps of 2%, the deadline increased for 100 ms and reached a maximum of 1000 ms (adults) or 1200 ms (children) with over 8% of time-outs. This was done to make sure that all participants produced a similar proportion of time outs (adults: $M = .02$, $SD = .01$; children: $M = .03$, $SD = .03$), and thereby had a similar opportunity to learn from feedback.

Data Recording

The experiment was controlled by the software E-Prime (PST Inc., Pittsburgh, PA). EEG and EOG activity were recorded continuously (Brain Amp DC, Brain Vision Recorder software) from 64 Ag/AgCl electrodes (10-10 system, American Electroencephalographic Society, 1994) using EasyCaps (Easycap GmbH, Herrsching, Germany). The left mastoid was used as reference and the data were re-referenced offline to averaged mastoids. EEG and EOG signals were filtered online from DC – 70 Hz and digitized at 500 Hz. Vertical and horizontal EOG was recorded from two electrode pairs placed on the infra- and supraorbital ridges of the right eye and on the outer canthi of the two eyes. Impedances were kept below 10 kΩ. To increase signal-to-noise ratio, a 30 Hz low-pass filter was applied. For the figures, a 20 Hz low-pass filter was applied.

Data Analysis

Accuracy Data. Responses faster than 140 ms were excluded from analysis (adults: 2.6%; children: 4.3%). We also excluded responses that exceeded the response deadlines. The analysis of variance (ANOVA) was based on accuracy (% correct.)[2] To analyze learning effects, we split each learning block into 4 quarters and averaged the quarters across the 5 blocks, yielding a total of 75 trials per subject, validity condition, and bin. To quantify the learning effects, we fitted a linear ($Y = b0+(b1 * t)$) function to each individual's learning curves, separately for the three validity conditions. The intercept (b0) and slope (b1 – or β –) parameters of the learning functions were then subjected to an ANOVA.

EEG Data. The EEG epochs (–200 to 600 ms) were averaged with respect to response and feedback onset. The response-locked EEG data were baseline corrected using a –200 to –50 ms pre-response baseline (see Hogan et al., 2005). For the feedback-locked EEG data, the average activity from –100 ms to feedback onset served as baseline. Prior to averaging, trials containing eye-movement artifacts or other artifacts were excluded from analysis using a threshold criterion (standard deviations greater than 30 μV within a sliding window of 200 ms). Remaining eye movements were corrected using a modified version of the linear regression approach (Gratton, Coles, & Donchin, 1983). The EEG data were processed using EEProbe software (ANT Software, Enschede, Netherlands). The response-locked ERPs were measured as mean amplitudes in a 0–100 ms time window post-response at electrode FCz. The Pe was measured as the mean amplitude between 300 and 400 ms at electrode Cz. The feedback-locked components were measured as mean amplitudes within a 100 ms time window centered on the peak of the FRN at FCz. For the peak-to-peak analyses, a 15 Hz low-pass filter was applied to obtain more reliable peak amplitude measures (see Frank, Woroch, & Curran, 2005). The ERN was defined as the peak-to-peak voltage difference between the most negative peak between –50 and 150 ms around the response and the preceding positive peak (see Frank et al., 2005). The FRN was defined as the difference between the most negative peak within 200 to 400 ms and the preceding positive peak.

[2]The accuracy rates in the 80% validity condition reflect the mean accuracy for the 80% valid trials of this condition. For the 20% invalid trials, mean accuracy was lower than chance (adults: $M = 34$, $SD = .11$; children: $M = 43$, $SD = .12$) because participants learned to respond to the dominant (but here incorrect) mapping.

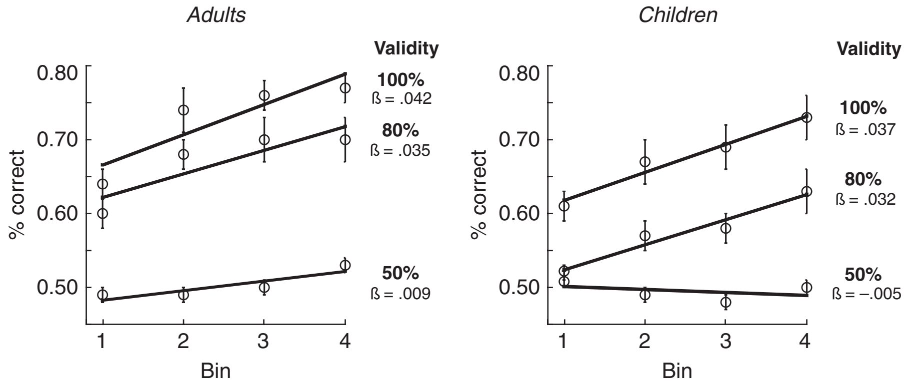

Figure 1. Accuracy learning curves for the three validity conditions, displayed separately for adults (left) and for children (right). The y-axis indicates accuracy in % correct; the x-axis shows the course of learning averaged into four bins.

Whenever necessary, the Geisser-Greenhouse correction was applied (Geisser & Greenhouse, 1958). In these cases the original *F*-value, the adjusted *p*-values, and the Epsilon values (ε) are reported. Bonferroni-corrections were applied when necessary (p-level $< .05$) and the corrected *p*-values are reported. Effects sizes (partial eta squared, η^2) are reported, which reflect the proportion of variance that is accounted for by the experimental manipulations (Cohen, 1973). As for the behavioral data, we averaged the ERPs into four bins. To quantify the learning effects, we fitted a linear ($Y = b0 + (b1 * t)$) function to each individual's learning curves separately for the three validity conditions. The intercept (b0) and slope (b1 – or β –) parameters of these learning functions were then subjected to an ANOVA. The statistical analyses were performed using SAS (SAS Institute Inc., Cary, NC) and SPSS (SPSS Inc., Chicago, IL).

Results

Accuracy Data

The accuracy data (% correct, see Figure 1 and Table 1) were analyzed using an ANOVA design with the factors Age group (adults, children), Validity (100%, 80%, 50%), and Bin (Bin1–Bin4). The analysis showed significant main effects of age group, $F(1,33) = 8.37$, $p < .007$, $\eta^2 = .20$, and validity, $F(2,66) = 102.02$, $p < .001$, $\varepsilon = .96$, $\eta^2 = .76$, as well as a significant two-way interaction between age group and validity, $F(2,66) = 4.59$, $p < .01$, $\varepsilon = .96$, $\eta^2 = .12$. Separate ANOVAs for the three validity conditions revealed that children performed worse than adults in the 80% validity condition ($p < .02$, $\eta^2 = .29$). No significant age differences were found for the 100% or 50% validity conditions ($ps > .12$). The analysis also showed a significant main

Table 1. *Mean Fit (R^2), Intercept (b0), and Slope (b1) Parameters of the Linear Learning Functions of the Accuracy Data, the Response-Locked Positivity, and the Feedback-Locked Positivity*

Age group	Validity	Bin	Accuracy				Response-locked positivity			Feedback-locked positivity		
			% correct	R^2	b0	b1 (β)	R^2	b0	b1 (β)	R^2	b0	b1 (β)
Adults	100%	1	.64 (.07)									
		2	.74 (.11)									
		3	.76 (.09)	.63 (.24)	.62 (.09)	.04 (.03)	.55 (.36)	2.3 (4.7)	1.0 (.80)	.60 (.33)	10.6 (3.9)	– 1.3 (.73)
		4	.77 (.10)									
	80%	1	.60 (.07)									
		2	.68 (.10)									
		3	.70 (.11)	.50 (.31)	.58 (.08)	.04 (.03)	.58 (.31)	1.4 (3.6)	0.8 (.86)	.36 (.29)	10.4 (4.2)	– 0.3 (1.1)
		4	.70 (.11)									
	50%	1	.49 (.04)									
		2	.49 (.05)									
		3	.50 (.04)	.26 (.31)	.48 (.04)	.01 (.02)	.36 (.30)	0.9 (4.2)	0.5 (.81)	.46 (.36)	11.5 (4.7)	0.1 (.10)
		4	.53 (.05)									
Children	100%	1	.61 (.09)									
		2	.67 (.11)									
		3	.69 (.13)	.52 (.36)	.58 (.10)	.04 (.03)	.40 (.28)	2.5 (4.9)	0.7 (1.3)	.30 (.26)	07.7 (4.8)	– 0.2 (1.2)
		4	.73 (.12)									
	80%	1	.52 (.06)									
		2	.57 (.07)									
		3	.58 (.09)	.55 (.29)	.50 (.08)	.03 (.04)	.53 (31)	– 0.0 (5.3)	1.1 (1.5)	.29 (.26)	08.6 (4.4)	0.1 (1.3)
		4	.63 (.11)									
	50%	1	.51 (.05)									
		2	.49 (.04)									
		3	.48 (.04)	.32 (.32)	.51 (.06)	– .01 (.02)	.28 (.27)	3.1 (4.3)	0.0 (1.6)	.44 (.33)	10.2 (6.3)	.01 (1.69)
		4	.50 (.04)									

Note: R^2 = mean fit parameter, b0 = intercept, b1 (β) = slope parameter of the learning function.

effect of bin, $F(3,99) = 30.83$, $p < .001$, $\varepsilon = .73$, $\eta^2 = .48$, and an interaction between validity and bin, $F(6,198) = 14.32$, $p < .001$, $\varepsilon = .70$, $\eta^2 = .30$. Separate analyses for the factor validity revealed significant learning effects for the 100% and the 80% validity conditions (ps $< .02$, η^2s $> .38$), but not for the 50% condition ($p < .08$).

To analyze age differences in the learning curves, we performed an ANOVA on the parameters of the learning functions (see Methods, see Table 1). The analysis of the intercepts revealed significant main effects of age group, $F(1,33) = 4.41$, $p < .04$, $\eta^2 = .12$, and validity, $F(2,66) = 18.03$, $p < .001$, $\varepsilon = .99$, $\eta^2 = .32$, and an interaction between age group and validity, $F(2,66) = 5.00$, $p < .01$, $\varepsilon = .99$, $\eta^2 = .09$. Contrasts for each of the levels of the factor validity revealed significant age differences in the intercepts only for the 80%–50% contrast ($p < .02$, $\eta^2 = .19$). The analysis of the slope parameters showed a significant effect of validity, $F(2,66) = 25{,}57$, $p < .001$, $\varepsilon = .78$, $\eta^2 = .44$. Contrasts for each of the levels of the factor validity revealed higher slope parameters for the 100% and the 80% condition than the 50% condition (ps $< .001$, η^2s $> .40$). No significant age differences in the slopes of the learning functions were obtained (ps $> .29$).[3] These findings show that age differences in the 80% condition were most pronounced at the beginning of learning (see Figure 1).

ERP Data

In the following section, we will first examine the ERP components to correct and incorrect responses (positive and negative feedback) using mean amplitude measures. Then, we will use peak-to-peak measurements for an additional quantification of the ERN and FRN. We decided for this procedure because the mean amplitude measures of these components are confounded by an overlapping positivity (see Figure 2). Learning-related effects in the ERPs will be examined by analyzing the parameters of the learning functions (see Methods).

Response-Locked ERPs

Figure 2 shows the ERPs to correct and incorrect responses in the three validity conditions for adults and children. In both age groups incorrect responses are followed by an ERN that increases the more valid the feedback. However, as also apparent from Figure 2, correct responses elicit a positivity (which will be termed response-locked positivity in the following) that also increases with feedback validity. Age differences are most pronounced in the 80% condition, in which children seem to differentiate less between correct and incorrect responses than adults. Following the ERN, an error positivity (Pe) can be observed that is particularly pronounced for children but also evident for adults.

The ANOVA for the mean amplitude measures of the response-locked components (measured 0–100 ms post-response) involved the factors Age group, Validity, Response type (correct, incorrect), and Bin. The analysis showed a significant effect of response type, $F(1,33) = 33.17$, $p < .001$, $\eta^2 = .50$, an interaction between response type and validity, $F(2,66) = 36.41$, $p < .001$, $\varepsilon = 83$, $\eta^2 = .52$, and a marginally significant interaction between age group, response type, and validity, $F(2,66) = 2.77$, $p = .08$, $\eta^2 = .08$. Separate analyses for the factors age group and validity revealed significant effects of response type for adults in the 100% and 80% conditions (ps $< .008$, η^2s $> .55$). In contrast, for children a significant effect of response type was obtained for the 100% condition ($p < .008$, $\eta^2 = .50$), but not for the 80% condition ($p < .20$), suggesting that children are impaired in differentiating correct from incorrect responses when feedback is less valid (see Figure 2). No significant effects were obtained for the 50% condition (ps $> .72$). Furthermore, we obtained a significant main effect of bin, $F(3,99) = 6.30$, $p < .002$, $\varepsilon = .79$, $\eta^2 = .16$, an interaction between response type and bin, $F(3,99) = 3.33$, $p < .03$, $\varepsilon = .85$, $\eta^2 = .09$ and a significant interaction between response type, validity, and bin, $F(6,198) = 4.50$, $p < .001$, $\varepsilon = .77$, $\eta^2 = .12$. Separate analyses for the factor response type showed a significant validity $\times$ bin interaction only for correct responses ($p < .03$, $\eta^2 = .08$). Post-hoc tests revealed an increase of the response-locked positivity with learning for the 100% and 80% validity conditions (ps $< .001$, η^2s $> .24$), but not for the 50% condition ($p < .63$) (see Figure 3a).

To analyze the learning-related effects in the response-locked positivity, we performed an ANOVA on the parameters of the learning functions (see Methods and Table 1). For the intercepts this analysis revealed a significant main effect of validity, $F(2,66) = 4.45$, $p < .02$, $\varepsilon = .97$, $\eta^2 = .11$, and an interaction between age group and validity, $F(2,66) = 4.17$, $p < .02$, $\varepsilon = .97$, $\eta^2 = .10$. Contrasts for each of the levels of the factor validity showed significant age differences only for the 80%–50% contrast, ($p < .02$, $\eta^2 = .20$). The analysis of the slope parameters revealed a significant main effect of validity, $F(2,66) = 7.21$, $p < .002$, $\varepsilon = .97$, $\eta^2 = .18$. Contrasts for each of the levels of the factor validity showed higher slope parameters for the 100% and 80% validity conditions than the 50% condition (ps $< .02$, η^2s $> .19$). No age differences in the slope parameters were obtained (ps $> .08$). Similar to the accuracy data, these findings show that age differences in the response-locked positivity in the 80% condition were most pronounced early during learning.

Peak-to-Peak Analysis of the ERN

The peak-to-peak measures of the ERN were analyzed using an ANOVA design with the factors Age group, Validity, and Bin. The analysis showed a significant main effect of age group, $F(1,33) = 6.20$, $p < .02$, $\eta^2 = .16$, suggesting that the ERN was larger for children ($M = -5.20$ μV, $SD = 1.38$) than for adults ($M = -3.33$ μV, $SD = 0.87$). Moreover, we obtained a significant main effect of validity, $F(2,66) = 19.29$, $p < .001$, $\varepsilon = .74$, $\eta^2 = .37$. Post-hoc contrasts showed that the ERN increased with feedback validity (ps $< .02$, η^2s $> .20$) (see Figure 2). However, no learning-related changes in the ERN were obtained (ps $> .13$).

[3]Because the linear function might underestimate age differences early during learning, we also fitted a second-order polynomial, which captures both the initially rapid learning rate and the later slower learning rate. For the early learning effects, the analysis revealed a significant main effect of age, $F(1,33) = 12.00$, $p < .002$, $\eta^2 = .27$, a significant effect of validity, $F(2,66) = 53.74$, $p < .001$, $\varepsilon = .99$, $\eta^2 = .58$, and a significant interaction between age group and validity, $F(2,66) = 5.77$, $p < .005$, $\varepsilon = .99$, $\eta^2 = .06$. Separate analyses for each of the levels of the factor validity revealed significant age differences only for the 80% validity condition ($p < .001$, $\eta^2 = .34$). For the late learning effects, the analysis revealed significant main effects of age group, $F(1,33) = 9.77$, $p < .004$, $\eta^2 = .22$, and validity, $F(2,66) = 22.65$, $p < .001$, $\varepsilon = .95$, $\eta^2 = .36$, and a significant interaction between age group and validity, $F(2,66) = 5.37$, $p < .008$, $\varepsilon = .95$, $\eta^2 = .09$. Separate analyses for each of the levels of the factor validity revealed a significant effect of age group only for the 80% condition ($p < .001$, $\eta^2 = .30$). Consistent with the linear learning functions, these findings suggest that age differences are most pronounced in the 80% validity condition at the beginning of learning.

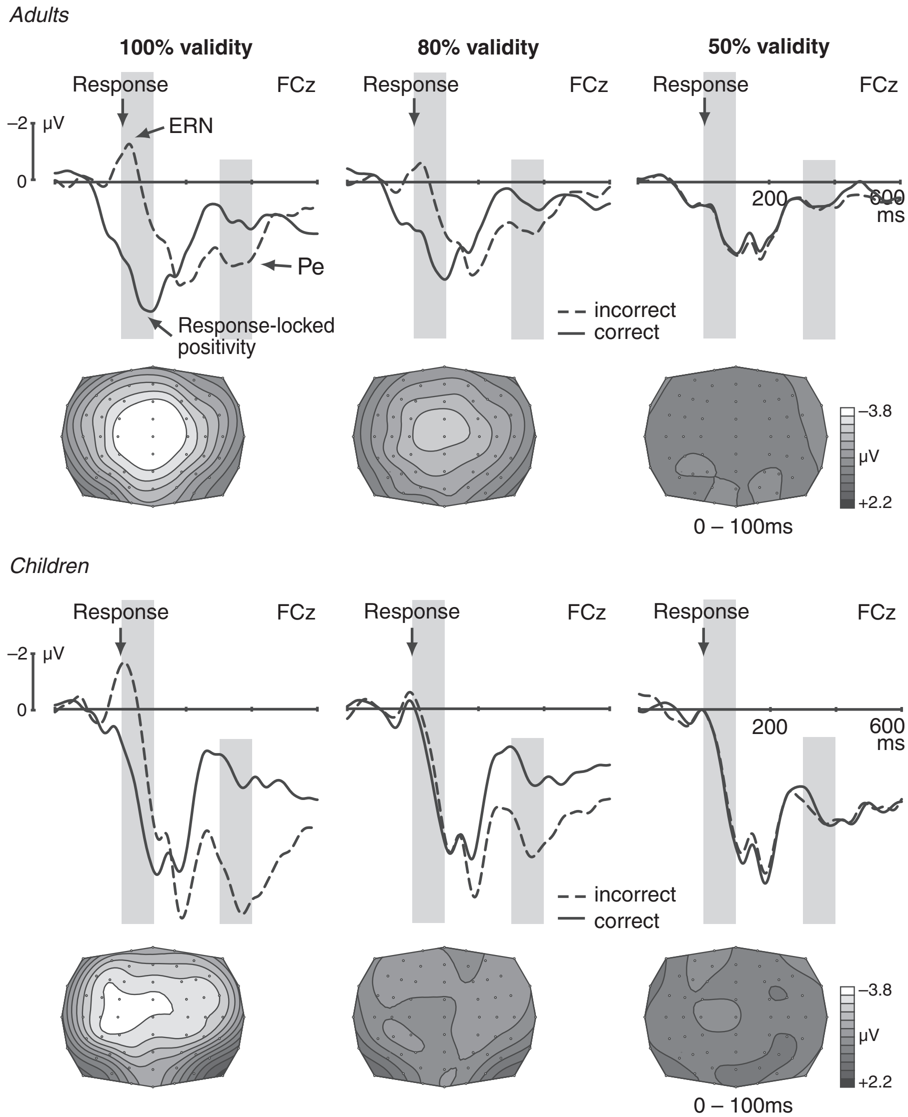

Figure 2. Response-locked grand average ERPs for the three validity conditions, displayed separately for correct responses (solid lines) and incorrect responses (dashed lines), for adults (top) and for children (bottom) at electrode FCz. Gray bars highlight the time windows used for statistical analysis. The topographic maps display the difference between the ERPs for correct and incorrect responses in a 100 ms post-response time window.

Error Positivity (Pe)

The mean amplitude measures of the Pe were analyzed using the same ANOVA as for the ERN. The analysis showed a larger Pe for children ($M = 7.27$ µV, $SD = 4.90$) than adults ($M = 1.44$ µV, $SD = 4.88$), $F(1,33) = 12.42$, $p < .001$, $\eta^2 = .27$. Moreover, we obtained a main effect of validity, $F(2,66) = 36.77$, $p < .001$, $\varepsilon = .73$, $\eta^2 = .53$. Post-hoc contrasts showed that the Pe increased with feedback validity (ps $< .02$, η^2s $> .33$), (see Figure 2). Interestingly, the analysis also showed a main effect of bin, $F(3,99) = 10.01$, $p < .001$, $\varepsilon = .90$, $\eta^2 = .23$, and an interaction between validity and bin, $F(6,198) = 2.48$, $p < .04$, $\varepsilon = .78$, $\eta^2 = .07$. Separate ANOVAs for the factor validity showed an increase of the Pe with learning for the 100% condition (ps $< .001$, η^2s $> .22$), but not for the 50% condition ($p < .67$).

Correlation Analysis

To investigate the relation between the response-locked ERPs and accuracy, we averaged the mean amplitude values for the response-locked positivity on correct trials and accuracy (in % correct) across the four bins and performed a correlation analysis using Pearson's correlation coefficients. As can be seen in Figure 3b, the amplitude of the response-locked positivity was significantly correlated with accuracy in the 100% validity condition, $r(35) = .57$, $p < .001$, and the 80% condition, $r(35) = .34$, $p < .05$, but not in the 50% condition, $r(35) = .19$, $p < .27$. Moreover, the correlations in the 100% condition were reliable for adults and children (rs $> .45$, ps $< .06$). Thus, better performance with learning is related to a larger amplitude of the response-locked positivity on correct trials. No reliable correlations for the ERN were observed (rs $< .24$, ps $> .17$).

Feedback-Locked ERPs

Figure 4 displays the ERPs to positive and negative feedback in the three validity conditions for children and adults. For both age groups a pronounced FRN can be observed for all validity conditions. Learning-related effects are more pronounced in the

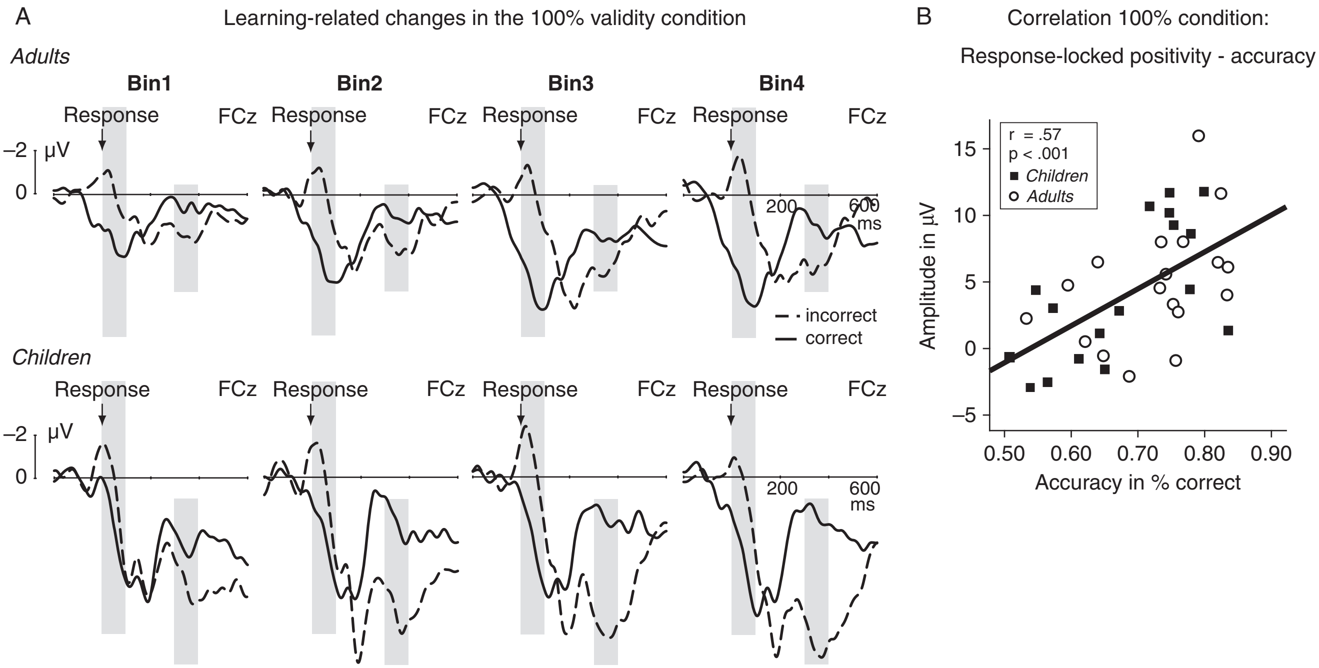

Figure 3. (A) Response-locked grand average ERPs for the 100% validity condition over the course of learning, displayed separately for correct (solid lines) and incorrect responses (dashed lines), for adults (top) and for children (bottom) at the electrode FCz. Gray bars highlight the time windows used for statistical analysis. (B) Scatter plot illustrating the correlation between accuracy (averaged across the 4 bins) on the x-axis (in % correct) and the mean amplitude of the response-locked positivity for correct trials on the y-axis (in μV) in the 100% validity condition. Children are displayed by squares and adults are displayed by circles.

ERP component for positive feedback (which will be termed feedback-locked positivity in the following) than the FRN (see Figure 5).

The ANOVA for the mean amplitude measures of the feedback-locked components included the factors Age group, Validity, Feedback type (positive, negative), and Bin. The analysis revealed a significant main effect of feedback type, $F(1,33) = 91.12$, $p < .001$, $\eta^2 = .73$, and an interaction between age group and feedback type, $F(1,33) = 6.28$, $p < .02$, $\eta^2 = .16$, which reflects the larger feedback effects for children than adults (see Figure 4). Separate analyses for the factor feedback type showed a marginally significant effect of age group for negative feedback ($p = .09$, $\eta^2 = .09$), but not for positive feedback ($p < .58$). Moreover, the analysis revealed a main effect of validity, $F(2,66) = 28.77$, $p < .001$, $\varepsilon = .79$, $\eta^2 = .47$, and an interaction between feedback type and validity, $F(2,66) = 4,73$, $p < .01$, $\varepsilon = .95$, $\eta^2 = .13$. Separate analyses for the factor validity showed that the difference in the ERPs to positive and negative feedback increased the more invalid the feedback ($ps < .001$, $\eta^2 s > .52$), (see Figure 4). The analysis also showed main effect of bin, $F(3,99) = 8.39$, $p < .001$, $\varepsilon = .92$, $\eta^2 = .20$, and an interaction between validity and bin, $F(6,198) = 2.89$, $p < .01$, $\varepsilon = .87$, $\eta^2 = .08$. Separate analyses for the three validity conditions showed a decrease of the feedback-locked positivity with learning for the 100% and the 80% validity conditions ($ps < .02$, $\eta^2 s > .10$), but not for the 50% condition ($p < .36$).

To quantify the learning-related effects in the feedback-locked positivity, we performed analyses on the parameters of the learning functions (see Methods). The analysis of the intercepts revealed a significant effect of validity, $F(2,66) = 3.38$, $p < .04$, $\varepsilon = .92$, $\eta^2 = .09$. Contrasts for each of the levels of the factor validity showed a larger intercept for the 100% than the 50% validity condition ($p < .02$, $\eta^2 = .18$). The analysis of the slope parameters showed a main effect of validity, $F(2,66) = 5.54$, $p < .009$, $\varepsilon = .87$, $\eta^2 = .14$, and a marginally significant interaction between age group and validity, $F(2,66) = 2.79$, $p = .08$, $\varepsilon = .87$, $\eta^2 = .08$. Contrasts for each of the levels of validity showed a higher slope parameter for the 100% than for the 80% and 50% validity conditions ($p < .02$, $\eta^2 = .20$). Separate analyses for the two age groups revealed significant effects of feedback validity on the slope parameters for adults ($ps < .001$, $\eta^2 = .47$) but not for children ($ps > .53$). These findings show that the feedback-locked positivity decreased with learning and that this decrease was less pronounced in children (see Table 1).

Peak-to-Peak Analysis of the FRN

The peak-to-peak measures of the FRN were subjected to an ANOVA involving the factors Age group, Validity, and Bin. The analysis only revealed a main effect of age group, $F(1,33) = 27.10$, $p < .001$, $\eta^2 = .45$, which reflects the larger FRN for children ($M = -14.88$ μV, $SD = 7.15$) than adults ($M = -6.13$ μV, $SD = 3.66$). However, no learning-related changes in the FRN were obtained ($ps > .12$).

Discussion

In the present study, we investigated developmental differences in error processing during reinforcement learning. We used a probabilistic learning task and examined the ERP correlates of internal and external error processing (ERN and FRN, respectively), comparing 10- to 12-year-old children with adults. The learning task involved three learning conditions in which we manipulated feedback validity (100%, 80%, and 50% validity).

The analysis of the accuracy data showed that in the 100% and 80% validity conditions children and adults learned from feedback, whereas in the 50% condition they performed at

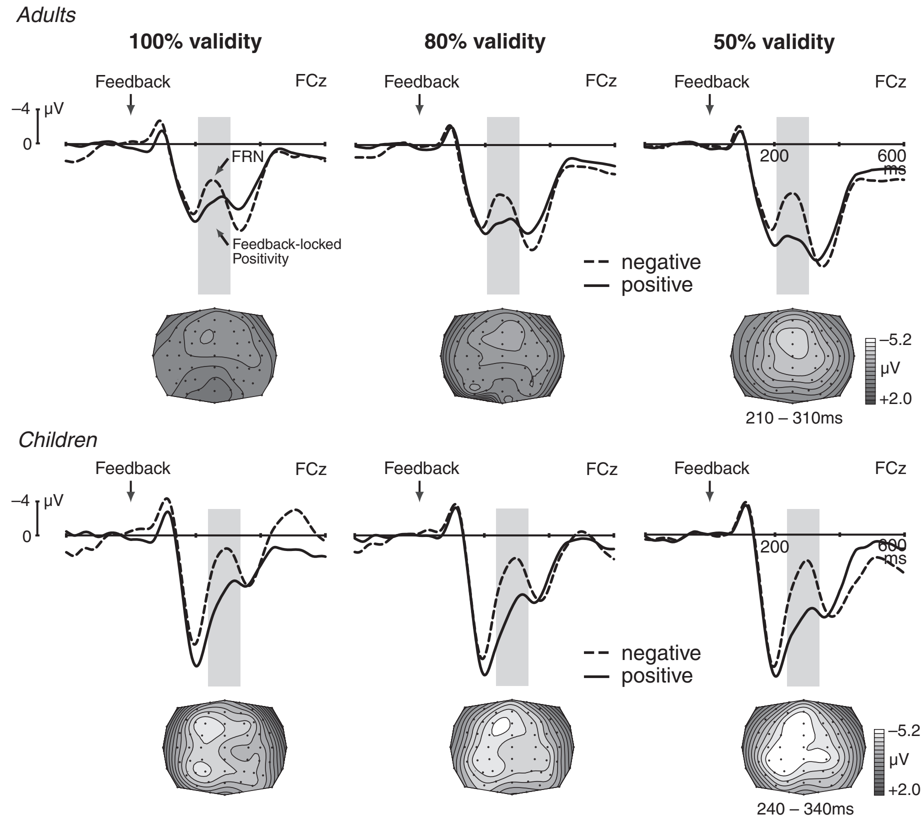

Figure 4. Feedback-locked grand average ERPs for the three validity conditions, displayed separately for positive (solid lines) and negative feedback (dashed lines), for adults (top) and for children (bottom) at the electrode FCz. Gray bars highlight the time windows that were used for statistical analysis. The topographic maps display the difference between the ERPs for positive and negative feedback in the time window of the FRN.

chance level. In the 100% validity condition, the two age groups showed comparable accuracy and similar learning rates. In contrast, in the 80% condition children showed reduced overall accuracy especially at the beginning of learning, suggesting that they are impaired in performance when invalid feedback interferes during learning (see Figure 1). These findings are in line with our expectations and with data from previous developmental and aging studies on learning and error processing (Eppinger et al., 2008; Stevenson & Weir, 1959). Taken together, the behavioral data show that children are as well able to learn as adults, if the feedback is fully valid and they receive enough time for responding. Yet, children are impaired, especially early during learning, when control requirements are enhanced due to interference by invalid feedback.

A very similar pattern of results was obtained in the response-locked ERPs. In the 100% validity condition, children and adults showed a pronounced ERN on incorrect trials, as well as a response-locked positivity on correct trials (see Figure 2). Thus, in the 100% condition in which children and adults learned comparably they also clearly differentiated between responses types. However, in contrast to adults, who also differentiated correct from incorrect responses in the 80% condition, children showed no significant effect of response type in this condition (see Figure 2). Consistent with the behavioral data, these findings suggest that children are not generally impaired in learning. Yet, they seem to be less able than adults to build up a stable representation of the correctness of a response when invalid feedback interferes with learning. The ability to represent interference-resistant response rules over time has been shown to be one of the key properties of the prefrontal cortex (Miller & Cohen, 2001). Hence, the present findings indicate that developmental differences in learning are most likely due to differences in the function of the prefrontal target areas of the dopamine system rather than differences in the learning signals themselves.

We did not obtain age differences in the amplitude of the ERN. This result stands in contrast to several recent findings, which showed a reduced ERN in children (Davies et al., 2004; Ladouceur et al., 2007; Santesso et al., 2006). However, in all of these studies smaller ERNs in children were paralleled by performance impairments, indicating that age differences in the ERN were confounded by performance differences between age groups. In line with previous findings (Eppinger et al., 2008; Hogan et al., 2005), the present results show that, when performance differences are absent between age groups, similar ERNs are observed for children and adults. This is not to say that any developmental differences in the ERN could be attributed to performance differences, and it may well be that the ERN follows a nonlinear trajectory when studied longitudinally, or across a larger age range than in the present study (Davies et al., 2004). However, in order to be able to interpret such results it seems necessary to control for performance differences between groups.

Furthermore, we did not find learning-related changes in the ERN, which seems inconsistent with the predictions of the R-L theory (Holroyd & Coles, 2002). Yet, it must be noted that the absence of learning effects in the ERN in the present study could result from the fact that the ERN bins were relatively large. That is, we might have missed rapid learning effects in the ERN at the beginning of learning. However, as could be seen in Figure 1,

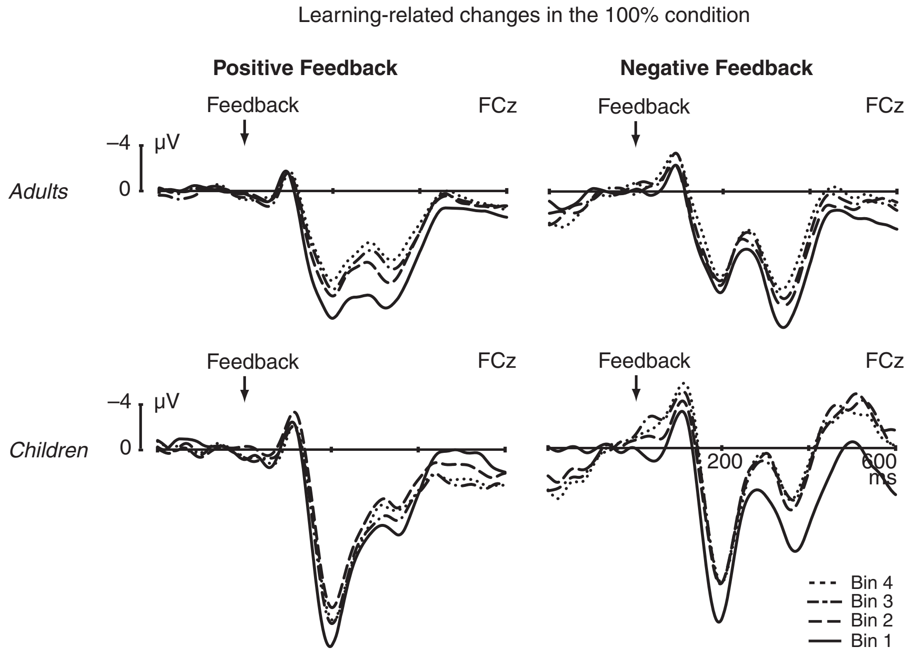

Figure 5. Feedback-locked grand average ERPs for positive and negative feedback trials in the 100% validity condition displayed separately for the four bins, for adults (top) and for children (bottom) at the electrode FCz.

there is considerable learning going on from bin 1 to bin 2 in both age groups, and according to the R-L theory one would have expected to see corresponding effects in the ERN. In contrast, we found an increase in the response-locked positivity with learning, which is consistent with previous findings that showed that learning-related changes are much more pronounced in the ERPs to correct than incorrect responses (Eppinger et al., 2008). The present findings further substantiate this view by showing a significant positive correlation between the amplitude of the response-locked positivity and accuracy (see Figure 3b). More research is needed to clarify to what extent the ERN co-varies with learning and whether it might reflect a rapid learning mechanism as opposed to a slower learning process as reflected in the response-locked positivity.

The ERN was followed by a centrally distributed positivity called Pe, which is assumed to either reflect error awareness (Nieuwenhuis et al., 2001), or the motivational significance of errors (Overbeek et al., 2005). Most of the studies so far have not found evidence for developmental changes in the Pe (Davies et al., 2004; Ladouceur et al., 2007; Santesso et al., 2006), whereas our data point to an increased Pe in children. This finding is consistent with results from a recent study that showed a larger Pe and more emotional reactivity to errors in adolescents than adults (Ladouceur et al., 2007). However, it must be noted that in the present study the error-related components were generally larger for children than adults, which makes it difficult to interpret the findings. Interestingly, the Pe also increased with feedback validity (see Figure 2). That is, the Pe was the larger the more participants were able to internally represent their errors, which is in line with the idea that the Pe reflects error awareness (Nieuwenhuis et al., 2001). However, since the present study does not involve a direct measure of error awareness, this finding seems equally consistent with the hypothesis that the Pe reflects the motivational significance of errors.

The analysis of the feedback-locked ERPs showed a larger FRN for children than adults, whereas no age differences were obtained in the ERPs for positive feedback (feedback-locked positivity). These findings suggest that children are more sensitive to negative feedback during learning than adults, whereas both age groups seem to be similarly affected by positive feedback.[4] Similar results have been obtained in a study on feedback processing in children with ADHD (van Meel et al., 2005). In the van Meel et al. (2005) study, ADHD children showed a larger FRN to negative feedback than controls, and the authors concluded that these children are more sensitive to unfavorable outcomes due to their dopaminergic dysfunction. Moreover, there is evidence from a developmental study using heart-rate measures that also suggests that children are more sensitive to negative feedback and less able to differentiate informative from uninformative feedback during learning (Crone et al., 2004).

We did not obtain reliable changes in the FRN with learning. In contrast, we found a learning-related modulation for positive feedback (feedback-locked positivity). The learning-related decrease of the feedback-locked positivity indicates that the more the participants were able to represent the correctness of the response, the less they relied on the feedback (see Eppinger et al., 2008). Hence, the present findings are in line with the idea that the feedback-locked positivity reflects the reduction of positive reward prediction errors with learning (Cohen, Elger, & Ranganath, 2007; Holroyd, Pakzad-Vaezi, & Krigolson, 2008). Interestingly, the learning-related decrease of the feedback-locked positivity was more pronounced for adults than children, suggesting that children are less able to disengage from external feedback during learning. Consistent with this result, the response-locked ERPs showed that children are impaired in build-

[4]The fact that we did not observed age differences in the feedback-locked positivity suggests that amplitudes were not generally different in children compared to adults. However, it could still be that anatomical differences in medial frontal cortex between age groups contributed to the larger FRN in children.

ing up an internal representation of the correct response, especially when feedback is less valid. An integration of these findings would suggest that children are less able than adults to internalize the information that is conveyed by the feedback. As a consequence, they have to rely more on external feedback than adults and show less pronounced learning-related changes in the feedback-locked positivity.

Conclusion

The present data show that children are as well able to learn as adults when feedback is always valid. However, they are impaired in performance and in building up a stable representation of the correctness of a response when invalid feedback interferes with learning. When performance levels are equated between age groups, children show similar ERN amplitudes as adults, which points to the importance of equating performance levels between age groups when interpreting age differences in the ERN. Our results further indicate that children show a larger feedback-related negativity (FRN) than adults, which may suggest that they are more sensitive to negative feedback during learning. Learning-related changes were only obtained in the ERPs to positive feedback and were less pronounced in children than adults, indicating that children are less able to disengage from external feedback during learning.

REFERENCES

American Electroencephalographic Society. (1994). Guideline thirteen: Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *11*, 111–113.

Biederman, J., & Faraone, S. V. (2005). Attention-deficit hyperactivity disorder. *Lancet*, *366*, 237–248.

Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement*, *33*, 107–112.

Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage*, *35*, 968–978.

Crone, E., Jennings, J. R., & van der Molen, M. W. (2004). Developmental change in feedback processing as reflected by phasic heart rate changes. *Developmental Psychology*, *40*, 1228–1238.

Davies, P. L., Segalowitz, S. J., & Gavin, W. J. (2004). Development of response-monitoring ERPs in 7- to 25-year-olds. *Developmental Neuropsychology*, *25*, 355–376.

Eppinger, B., Kray, J., Mock, B., & Mecklinger, A. (2008). Better or worse than expected? Aging, learning, and the ERN. *Neuropsychologia*, *46*, 521–539.

Falkenstein, M., Hohnsbein, J., & Hoormann, J. (1995). Event-related potential correlates of errors in reaction tasks. *Electroencephalography and Clinical Neurophysiology*, *44*, 287–296.

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, *47*, 495–501.

Gehring, J. W., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385–390.

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*, 2279–2282.

Geisser, S., & Greenhouse, S. W. (1958). An extension of box‘s results on the use of the F-distribution in multivariate analysis. *Annals of Mathematical Statistics*, *29*, 885–891.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for offline removal of ocular artefact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484.

Hogan, A. M., Vargha-Khadem, F., Kirkham, F. J., & Baldeweg, T. (2005). Maturation of action monitoring from adolescence to adulthood: An ERP study. *Developmental Science*, *8*, 525–534.

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679–709.

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G. H., et al. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, *7*, 1–2.

Holroyd, C. B., Baker, T. E., Kerns, K. A., & Mueller, U. (2008). Electrophysiological evidence of atypical motivation and reward processing in children with attention-deficit hyperactivity disorder. *Neuropsychologia*, *46*, 2234–2242.

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, *45*, 688–697.

Johnson, M. H., & Munakata, Y. (2005). Processes of change in brain and cognitive development. *Trends in Cognitive Sciences*, *9*, 152–158.

Ladouceur, C. D., Dahl, R. E., & Carter, C. S. (2007). Development of action monitoring through adolescence into adulthood: ERP and source localization. *Developmental Science*, *10*, 874–891.

Liotti, M., Pliszka, S. R., Perez, R., Kothmann, D., & Woldorff, M. G. (2005). Abnormal brain activity related to performance monitoring and error detection in children with ADHD. *Cortex*, *41*, 377–388.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.

Miltner, W. H., Braun, C. H., & Coles, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a ‘generic’ neural system for error detection. *Journal of Cognitive Neuroscience*, *9*, 788–798.

Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioral control. *Nature*, *431*, 760–767.

Müller, S. V., Möller, J., Rodriguez-Fornells, A., & Münte, T. F. (2005). Brain potentials related to self-generated and external information used for performance monitoring. *Clinical Neurophysiology*, *116*, 63–74.

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology*, *38*, 752–760.

Nieuwenhuis, S., Yeung, N., Holroyd, C. B., Schurger, A., & Cohen, J. D. (2004). Sensitivity of electrophysiological activity from medial frontal cortex to utilitarian and performance feedback. *Cerebral Cortex*, *14*, 741–747.

Offenbach, S. I. (1964). Studies of children’s probability learning behavior: I. Effect of reward and punishment at two age levels. *Child Development*, *35*, 709–715.

Overbeek, T. J. M., Nieuwenhuis, S., & Ridderinkhof, K. R. (2005). Dissociable components of error processing: On the functional significance of the Pe vis-à-vis the ERN/Ne. *Journal of Psychophysiology*, *19*, 319–329.

Santesso, D. L., Segalowitz, S. J., & Schmidt, L. A. (2006). Error-related electrocortical responses in 10-year-old children and young adults. *Developmental Science*, *9*, 473–481.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, *36*, 241–263.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, *6*, 174–215.

Stevenson, H. W., & Weir, M. W. (1959). Variables affecting children's performance in a probability learning task. *Journal of Experimental Psychology*, *57*, 403–412.

van Meel, C. S., Oosterlaan, J., Heslenfeld, D. J., & Seargent, J. A. (2005). Telling good from bad news: ADHD differentially affects the processing of positive and negative feedback during guessing. *Neuropsychologia*, *43*, 1946–1954.

van Veen, V., & Carter, C. S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, *14*, 593–602.

(Received May 2, 2008; Accepted November 10, 2008)